

Common sense in the environment

A scientific panel of the US National Research Council has shown that too little of what passes for environmental clean-up meets the test of scientific rigor, while a US court overturns a ban on asbestos.

EVER since Rachel Carson launched the environmental movement with her compelling warning that the world might one day awaken to a "silent spring", the United States has taken the threat of environmental ruin with increasing seriousness. Laws have been written to regulate waste disposal and pesticide use. Agencies such as the Environmental Protection Agency (EPA) and the National Institute of Environmental Health Sciences have been created. And the courts have played their part in ordering financial redress to people who seem to have been injured by environmental hazards, despite lack of supporting evidence.

It comes as something of a surprise, then, to hear no less conservative a voice than that of the US National Academy of Sciences claim that the United States may be wasting billions of dollars because, in its zeal to clean up the environment, the government has failed to conduct studies that would provide a sound scientific underpinning to judgments about which of the many potential toxins that surround us constitute a genuine and serious threat to human health. But that is just what the academy has to say in a new National Research Council report, "Environmental Epidemiology: Public Health and Hazardous Wastes". Under legislation known colloquially as Superfund — a federal bank account to pay for waste clean-up — 1,200 toxic waste dumps have been designated as the most hazardous of more than an estimated 400,000 sites (large and small) throughout the United States.

Health data needed

Now the academy challenges the criteria for designating the 1,200 sites as more dangerous than others, saying "A decade after implementation of Superfund, and despite congressional efforts to redirect the program, substantial public health concerns remain, and critical information on the distribution of exposure and health effects associated with hazardous-waste sites is still lacking." Decisions about clean-up have been made on political rather than scientific grounds, such that "less than 1 percent of the \$4,200 million spent *each year* (emphasis added) on hazardous-waste sites in the U.S. has been used to evaluate health risks at listed Superfund sites."

This is a stinging indictment of the policies of a host of federal agencies, including not only the EPA but also the departments of defence and energy, which spend another

\$6,000 million or so a year on environmental clean-up. But the issue is not just one of money wasted. The point is that decisions to spend money on ugly but perhaps less dangerous sites may inadvertently leave the public vulnerable to more serious health threats at sites that have not made the list of 1,200.

An obvious response to the academy's assertion that clean-up decisions are made on the grounds of engineering feasibility rather than long-term public health considerations is that, in the political arena, one cannot wait for perfect data before taking action. But that begs the question, particularly in this case when literally hundreds of millions of dollars are available and the issue is their proper distribution, which ought to include gathering scientific data.

Groundwater contamination

Groundwater contamination is but one example. Data show that 50 per cent of the US population depends on groundwater. In California, the figure is 70 per cent. Furthermore, it has been ascertained that at certain sites that are not top priority, groundwater contamination is substantial and (equally important) localized. Thus, direct epidemiological studies of population exposure are possible. For instance, researchers could use biological markers such as the concentrations of contaminants in blood and urine, or symptoms such as headache, fatigue or (even more important) birth defects, to monitor the health effects of toxic exposure.

It is the academy's position, based on extensive review of the literature, that although such markers are frequently employed to ascertain occupational exposure, they have not been adequately adopted for assessment of broad environmental exposure. This leads to a real danger — that if (as anticipated) groundwater contamination spreads, becoming almost universal in the United States, it will in future not be possible to assess the health effects. In short, a so-called window-of-opportunity exists for environmental epidemiology that must be seized now or forever lost.

Doing the right experiments matters. The United States after all is not the only country whose residents are exposed to toxic materials in the environment, but it is one where the scope of the problem and the resources for doing broadly useful scientific studies come together. Good environmental epidemiology would be applicable to deal-



ing with future crises like those at Seveso, Italy and Bophal, India. It is not enough to invest gigantic sums in environmental clean-up or in compensation to victims if the true extent of their injury remains uncertain.

Arguments for better epidemiological studies are all the more compelling because the research foundation in the field already exists, including protocols for identifying 'sentinel' events (for instance, an unusual incidence of congenital cardiac anomalies often follows exposure to trichloroethylene in water). Yet in most cases, assessment of risk at hazardous-waste sites relies too heavily on inadequately validated modeling rather than actual measurement of nearby residents.

Indeed, the need to separate the health effects of toxic exposure from real but excessive public anxiety is also now evident in decisions handed down by the federal courts. Last month, a US appeals court overturned an EPA ban on asbestos, arguing that the environmental agency failed to weigh financial costs against public health benefits when it banned the use of asbestos in products including insulators, fireproof fabric, automobile brake linings and various building materials. The Toxic Substances Control Act (TSCA) of 1976, on which EPA based its ban, requires that agencies conduct a risk/benefit analysis of their actions. The court said, "...we do note that the EPA, in its zeal to ban any and all asbestos products, basically ignored the cost side of the TSCA equation."

By EPA's own figures, a total asbestos ban would save the lives of 202 people during a span of 13 years, at a cost ranging from \$76 million to \$106 million per life. No one disagrees with data that show inhaling asbestos fibres to be pathogenic to the lungs, but new scientific studies suggest that asbestos in solid form is not the hazard that free-floating asbestos fibre is. Furthermore, the court said, "Considering that many of the substances that the EPA itself concedes will be used in the place of asbestos have known carcinogenic effects, the EPA not only cannot assure this court that it has taken the least burdensome alternative, but cannot even prove that its regulations will increase workplace safety."

What we are seeing is an attempt by some researchers and judges to infuse the environmental movement with greater scientific rigor and common sense. There was a time when such thinking would have been regarded as heretical. But it is high time that rigor and common sense made their way into the debate about protecting the environment, which will in the long run be good for the environment and the public health. □

Recipe for prosperity?

US should convert emphasis on military technology to include commercial technology.

Does it matter that only 25 per cent of telephones in the land of Alexander Graham Bell are made in the United States? The decline in US commercial technology has been long coming and often even arrogantly dismissed as

insignificant because US companies retained world leadership at the cutting edge of high technology even as they lost market share in mass produced products. But can the US economy thrive on the strength of the few areas in which US companies still compete successfully in the world market — chemicals, pharmaceuticals and aircraft among them?

According to a new analysis by the Carnegie Commission on Science, Technology and Government (*Nature* **353**, 198; 19 September 1991), the United States must commit itself to a strategy for the development of its technology base, commercial technology included. The challenge is particularly urgent in the light of the sudden, unanticipated demise of the Soviet Union as a military threat and the consequent changes in spending priorities at the Pentagon. Carnegie says that ever since the Second World War, the US has invested heavily in military technology, relying on spinoffs from Department of Defense research for innovation in private industry and manufacturing. But now, the Carnegie commission argues with good evidence, "defense-supported technology lags rather than leads the marketplace in many areas."

Republican administrations have followed fashion elsewhere in the past decade in staying shy of anything that might come close to a national technology policy, holding that a free market takes care of everything. True, the administration has paid lip service to the cause of competitiveness by setting up committees throughout the government, but these good intentions have been fitful and unfocussed; no agency nor single office in the White House is in charge. The Carnegie commission makes a number of specific suggestions. It urges President George Bush to vest more authority in the Office of Science and Technology Policy, now headed by White House science adviser D. Allan Bromley (whose concern about technology policy is well known). The commission also calls for the transformation of the intellectually elite Defense Advanced Research Projects Agency (DARPA) into the National Advanced Research Projects Agency (NARPA), thus shifting the focus from military technology to what is called 'dual-use' technology.

These proposals would probably do very little harm, but they do not get to the root of the problem. Indeed, they have a quaintly old-fashioned air, reminiscent of the way in which European governments (notably the British and the French) have in the past sought to keep Japan's manufacturers at bay by sponsoring industrially relevant research. Most of these schemes have merely demonstrated that civil servants, however steeped in defence technology, are unskilled at telling what the commercial market needs. Yet present bickering over the European Community's support for manufacturing technology shows that the truth has not sunk home even in Europe, which abounds with past failures in the public sponsorship of industrial research. A more coherent policy from the US government might, as the Carnegie suggests, be useful but it is not sufficient. The private sector must take the lead and show some initiative of its own. □

French blood scandal echoes across Europe

- Three French officials will stand trial
- Transfusion services elsewhere are watching

London

THE forthcoming trials of three French health officials are causing concern among blood transfusion services in other European countries, where similar charges might be brought. Jacques Roux, former director general of the health ministry, and Robert Netter, former head of the French national health laboratory, were charged last week with failing to prevent the distribution of HIV-infected blood-clotting factors to French haemophiliacs in 1985. Michel Garretta, former head of the French national blood transfusion centre (CNTS) in Paris, is accused of knowingly supplying the infected clotting factors.

The charges have destroyed the morale of the country's transfusion service. But blood transfusion experts warn that the shockwave from the trial may extend beyond France. Because France was by no means the last European country to begin heat-treating clotting factors supplied to haemophiliacs (which deactivates any HIV present), the action against the Frenchmen sets a worrisome precedent for their counterparts elsewhere in Europe.

"The problem is not a French problem," insists Bahman Habibi, acting medical and scientific director of CNTS. Johannes Leikola of the Finnish Red Cross blood transfusion service agrees that, seen from a European perspective, the French charges seem "grossly unfair".

The charges were levelled first by a militant haemophiliac group, the Association of the Polytransfused. Whether other European groups representing haemophiliacs, many of which are still seeking compensation for their infected members, will now follow the French example and seek to bring individual officials to trial is still unclear. Cees Smit, coordinator of the Dutch Haemophilia Society, says that no such action is likely in the Netherlands, where the blood transfusion service began consulting with his group about the possibility of HIV infection in late 1982. But he says that in several southern European countries, there was "a large denial of the HIV problem" in the early 1980s, so haemophiliacs there may have similar complaints to those in France.

Last week's announcement that the three French officials will stand trial follows a two-year investigation by the French justice department. The cases are expected to focus on the period from February to October 1985. In February 1985, a study published in the *Lancet*, co-authored by Pasteur Institute AIDS re-

searcher Luc Montagnier, showed that heat treatment of blood-clotting factors could protect haemophiliacs from HIV infection. By October, all clotting factors released onto the market in France were heat treated.

The precise number of haemophiliacs who became infected during this time is still unknown, but in an interview earlier this year with the French consumer magazine *Que Stoisir*, Jean-Pierre Allain, who in 1985 was responsible for blood plasma derivatives research at CNTS, said that CNTS officials had estimated that 30 to 60 French haemophiliacs could become infected each month.

Bruno de Langre, president of the French Haemophilia Association, says that CNTS and the health ministry should have moved more quickly to ensure that all French-produced clotting factors were heat treated, and to ensure that imported clotting factors came only from companies that already used heat treatment.

De Langre points out that some commercial US manufacturers of clotting factors began heat treating their products as early as 1983. But Finland's Leikola, who is also vice-president of the European Plasma Fractionation Association (EPFA), says that this was done to reduce the transmission of hepatitis C, rather than HIV. In Europe, where hepatitis C was a smaller problem, heat treatment was not thought necessary.

Leikola believes the French introduction of heat treatment for blood-clotting factors cannot be described as sluggish, at least when compared with other blood transfusion services in Europe. Last year, he surveyed the seven leading non-profit laboratories responsible for producing factor VIII, the main clotting factor supplied to haemophiliacs. The laboratories included the main suppliers in Australia, Finland, France, the Netherlands, Scotland (whose blood transfusion service is run separately from that in England and Wales) and Switzerland. Heat-treated factor VIII from the Commonwealth Serum Laboratory in Australia was the first to become available — in November 1984 — but it was not until October 1986 that all of the European laboratories he surveyed were marketing heat-treated factor VIII.

Pim van Aacken of the Dutch Red Cross blood transfusion service, who serves as EPFA president, adds that in the mid-1980s it was still uncertain whether heat-treated clotting factors were safe. Many physicians feared that heat treat-

ment would cause the factors to denature and that this would cause haemophiliacs to produce antibodies against factor VIII, he says. "People forget the uncertainty that existed from 1983 to 1985."

Whatever the implications for the rest of Europe, the charges against Garretta, Netter and Roux have had a shattering effect on the French blood transfusion service. "People are demoralized — they've lost confidence in their jobs," says Habibi of CNTS. Habibi also believes that attention has been focused unfairly on the role of his centre. Although CNTS does control the import of all clotting factors into France, in 1985 it was responsible for less than half of the French production of factor VIII — the rest coming from six other non-profit making laboratories under the aegis of the health ministry. Habibi says that, from July 1985, all clotting factors distributed from CNTS were heat treated.

The turmoil in the French transfusion service is set to worsen. Separate accusations that the French government delayed the licensing of the HIV-diagnostic blood test developed by the US company Abbott so that a competitor test from the French company Diagnostic Pasteur could be released first (see *Nature* 353, 197; 19 September 1991) may lead to further charges. The investigation is expected to question the actions of senior figures in the government of the day.

Peter Aldhous

NATIONAL INSTITUTES OF HEALTH — Score wars

Washington

IN an attempt to fund the US National Institutes of Health fully in the midst of a budget crunch, Congress is resorting to some accounting sleight-of-hand that may delay as many as 1,800 grants for several months and force researchers to turn to their institutions for help. Following a proposal by the Administration in its original 1992 budget request, Congress is planning to delay up to \$600 million of NIH's budget until 30 September 1992 — the end of the fiscal year. The point is to 'score' the money in the 1993 budget, so that Congress can spend more than it was allocated for NIH in 1992. But university research advocates are understandably suspicious of such legerdemain.

Dave Moore of the American Association of Medical Colleges estimates that research universities may be forced to spend \$1 million each in keeping laboratories open while researchers wait two months for their promised funding to start. About one-third of NIH's 5,785 grants are expected to be affected. NIH budget policy chief Rita Koch says that although the agency is considering several plans that may minimize the impact of the scheme, "every option that we're looking at involves some delay of grants." C.A.

Japan promotes nuclear power

Tokyo

DESPITE a major accident earlier this year at the Mihama nuclear power plant that shook public confidence in nuclear power, the Japanese government intends to press ahead with ambitious plans to expand the use of atomic energy, according to the latest annual report on atomic energy released by the Science and Technology Agency (STA) last week. But growing public opposition to the siting of nuclear plants may thwart the government's plans.

The annual report devotes several pages to the Mihama accident in February, which was the first time an emergency cooling system had come into play in Japan to prevent meltdown of a reactor core (*Nature* 349, 557; February 1991). The report argues that the accident, caused by a rupture in the primary cooling system of the reactor, demonstrated that the "engineering safety" of Japanese nuclear power plants "has been confirmed" because the emergency cooling system functioned properly and there were no harmful effects on the environment.

But the report admits that Mihama has aroused feelings of "anxiety" in the general public. And it calls for promoting a "correct understanding" of the need for nuclear power with lectures, telephone answering services, personal computer communication links, visits to plants and the loaning of simple radiation-measuring equipment to the public. The report also says it is important to give actual examples

that show how the construction of nuclear power plants contributes to the development of local communities in the vicinity.

The government has a powerful array of subsidies to offer local governments that accept nuclear power plants (see below). But, despite the vast sums of money local communities can receive, very few are now prepared to accept new plants.

The government's plans call for a doubling in the number of nuclear plants to about 80 by the year 2010 to supply 43 per cent of Japan's electricity needs, compared with the present 40 commercial plants that provide about 26 per cent. But of the approximately 40 new 1-megawatt plants needed to meet the government's target, sites have only been found for 12 and most of these are on the grounds of existing nuclear power plants.

Hideaki Tsuzuku, deputy director of the Office of Atomic Energy Policy Research of STA's Atomic Energy Bureau, admits that the government is "under very trying circumstances in trying to obtain in a very credible manner the targets we have set."

Another problem facing the government is the need to introduce the use of highly toxic and dangerous plutonium to Japan's ordinary nuclear power plants over the next decade. Starting in the second half of next year, Japan will import plutonium from France and the United Kingdom that has been reprocessed from spent Japanese nuclear fuel. Japan is under contract to ship about 30 tons of plutonium from

Europe by 2010, Tsuzuku says.

Originally the government planned to use the plutonium in fast breeder reactors. But, with development of commercial fast breeders delayed by decades, most of the plutonium will have to be used in ordinary reactors by mixing it with uranium fuel because it would be against the nuclear non-proliferation treaty, to which Japan is a party, to store vast quantities of weapons-grade plutonium.

However, it is very much in doubt whether local communities will accept the use of plutonium in nearby plants. At the very least, the government will have to come up with even larger subsidies to promote acceptance. **David Swinbanks**

X-RAY ASTRONOMY

Ginga to fall soon

Tokyo

JAPAN'S X-ray astronomy satellite Ginga is expected to burn up in the Earth's upper atmosphere sometime this week after producing a wealth of data for the world's astrophysicists during the past five years.

Ginga was launched by Japan's Institute of Space and Astronautical Science (ISAS) in February 1987 and carried X-ray detectors developed in collaboration with Leicester University in England and a US-built gamma-ray burst detector (*Nature* 325, 567; 1987). It cost only about \$50 million to build and launch — a fraction of the cost of US and European space satellites.

Within months of launch the satellite made scientific headlines by detecting X-rays from the supernova in the Large Magellanic Cloud. This was followed by a string of other achievements, including discovery of two very bright galactic transients that are strong candidates as black holes; detection of cyclotron absorption lines in gamma-ray burst spectra thought to come from a highly magnetized neutron star; discovery of five new X-ray pulsars; detailed studies of the quasi-periodic oscillations of a neutron star; detection of dramatic rises in luminosity of quasars; and sensitive studies of cosmic X-ray background.

Approximately 100 scientists have made use of Ginga. The satellite has produced 17 PhD theses, including three each from the United States and United Kingdom, and several more are in preparation.

At the time of *Nature* going to press, ISAS scientists predicted that Ginga is most likely to plunge into the atmosphere on Friday 1 November, but its demise could come at any time between 31 October and 3 November. ISAS will continue to collect data right up to the last minute. But the satellite's bumpy ride through the upper atmosphere is making it difficult to point the satellite's detectors in a fixed direction during its final hours.

David Swinbanks

The persuasive power of yen

Tokyo

THE Japanese government provides lavish subsidies to local communities that accept nuclear power. Even so, finding new sites for nuclear plants is proving increasingly difficult (see above).

Under a law introduced in 1974 after the Middle East oil crisis, local governments that accept electric power plants or nuclear power-related facilities are entitled to a variety of central government subsidies. The subsidies are particularly large for nuclear power plants.

This fiscal year, ¥178,700 million (\$1,400 million) has been set aside under the law, 9.5 per cent more than last year. The funds come from a tax on electricity use that currently runs at ¥445 per thousand kilowatt-hours. The funds are not subject to the strict ceilings of other government budgets, and use of electricity is growing rapidly in Japan.

About half of the funds (¥87,100 million) are in the form of subsidies to build new local community facilities. The size of subsidy depends on the kilowatt

output and type of power plant. Nuclear power plants get by far the largest subsidies.

The other half goes mainly to subsidies for local industrial development (¥16,200 million), promotion of public acceptance of electric (in particular, nuclear) power plants (¥12,900 million), and subsidies and subcontracts for testing and checking the safety and reliability of nuclear power plants (¥36,200 million).

In the case of the little village of Rokkasho in northern Japan, which has accepted three commercial plants for the reprocessing and enrichment of nuclear fuel and storage of nuclear waste, the government will provide ¥18,000 million (\$140 million) in subsidies over ten years. The money, which amounts to more than \$10,000 for every man, woman and child in the village, will be used for 96 village projects, including radar facilities for fishermen, a fertilizer plant for farmers, a sports amenity centre and a ski resort. **D.S.**

Nuclear science faces cuts

Washington

THE US Department of Energy (DOE) plans to cut spending on nuclear science by 10 per cent in fiscal year 1993, a move that will probably force the closing of at least one major nuclear science facility.

William Happer, director of DOE's Office of Energy Research, met with members of the Nuclear Science Advisory Committee in East Lansing, Michigan, last week and asked them to recommend how to rearrange spending priorities in the face of such a cut. It won't be easy, said Peter Paul, a physics professor at the State University of New York at Stony Brook and chairman of the advisory committee. "Nobody really wants to accept such a cut," he said. "We will have to shut down one of our two main projects" — either the Los Alamos Meson Physics Facility or the Relativistic Heavy Ion Collider (RHIC), now under construction.

Faced with increasing environmental costs and an overall spending cap imposed by the Budget Enforcement Act of 1990, the DOE has been forced to cut back sharply on its planned spending for basic research (see *Nature* 353, 287; 26 September 1991). The decreases have already forced the fusion community to cancel its Burning Plasma Experiment, which was

planned as the successor to the Tokamak Fusion Test Reactor (*Nature* 353, 372; 3 October 1991), and the High Energy Physics Advisory Panel met 28-29 October to discuss how to deal with smaller budgets for that field.

Although Paul would not give precise details on what the nuclear science panel will recommend to Happer, he did say that the committee "reaffirmed the high scientific value of RHIC". This seems to imply that if something has to go, it will probably be the meson facility at Los Alamos National Laboratory. But closing that facility will not be enough, Paul said — several smaller programmes will also have to be closed if the nuclear science budget is to be cut by 10 per cent.

Happer also asked the committee to make recommendations for the four years after 1993, assuming three different scenarios: spending that remains fixed at the 1993 level without taking inflation into account; budgets that stay even in real, inflation-adjusted terms; and budgets that have 2 to 3 per cent real growth. The committee's recommendations for fiscal year 1993 should be ready soon, Paul said, but it will take longer to decide on how to set priorities for the longer-range budgets.

Robert Pool

Ozone report puts US policy in question

London

ACCORDING to US environmentalists, last week's United Nations Environment Programme (UNEP) report on the state of the ozone layer (see *Nature* 353, 688; 24 October 1991) not only demands urgent action to further reduce emissions of ozone-depleting chemicals, but also undermines the US Administration's policy on climate change.

The United States has so far refused to set a target date by which US emissions of carbon dioxide will be stabilized, even though other nations have set dates. The Bush Administration has instead argued that total US emissions of greenhouse gases will be no higher in 2000 than they were in 1987 (see *Nature* 348, 4; 1990). Although US emissions of carbon dioxide are forecast to rise, this will be offset by reductions in the emissions of other greenhouse gases, according to a 1990 report from the Environmental Protection Agency.

The EPA prediction is based on the relative warming effects ascribed to different greenhouse gases by the Intergovernmental Panel on Climate Change (IPCC) in its 1990 report. But according to the new UNEP ozone report, it seems that the warming caused by chlorofluorocarbons (CFCs) may be less than was thought by the climate panel. UNEP's ozone monitoring experts now say that CFCs can have both a warming and a cooling effect (the latter occurs because ozone depletion causes cooling of the lower stratosphere, which then radiates less heat down into the troposphere). Colin Johnson, from AEA Environment and Energy in Oxfordshire, and a member of the UNEP group, says the warming and cooling effects "seem rather similar in magnitude". When the UNEP report was completed earlier this month, he says, several atmospheric scientists suggested that this similarity could explain why the observed increase in global temperatures has so far lagged behind theoretical predictions.

If CFCs cause less warming than was thought, then the US Administration's claim that reductions in emissions of CFCs will offset the warming caused by increased US production of carbon dioxide may no longer hold true. Dan Lashof from the Natural Resources Defense Council is cautious — the cooling effects of CFCs have been shown only in some regions and at particular times of year, he says — but he argues that the new data support the line taken by the European Communities (EC) at last month's climate convention negotiations in Nairobi. At that meeting, EC negotiators rejected the US position that any controls on greenhouse gases should take into account all emissions, not just carbon dioxide.

Peter Aldous

New quarters for the Superconducting Super Collider



Its future is still very much up in the air, but the Superconducting Super Collider (SSC) now at least has one foot on the ground. This month the laboratory completed its first permanent building: a 100,000-square-foot facility in which researchers will engineer and construct some of the superconducting magnets that will guide and focus particle around the planned 54-mile underground ring.

Olympic row over sex testing

Washington

A PROMINENT Spanish geneticist has refused to participate in a sex testing trial of women athletes participating in the 1992 Olympics in Barcelona, arguing that the simplicity of new testing methods, combined with their inherent fallibility, could lead to increased misdiagnosis of women athletes as men.

Xavier Estivill, of the Cancer Research Institute at the Turan Reinal Hospital in Barcelona, says he has turned down a request by the International Olympic Committee to conduct a trial sex test using gene amplification technology on some 300 of the 3,000 women who will participate in the 1992 Olympics.



The technology based on polymerase chain reaction (PCR) analysis is, he says, so easy to use that sports-related genetic sex-testing may become commonplace, despite continuing scientific debate about what defines a woman and a dearth of genetic counsellors trained to explain to young female athletes what it means, for example, to have some Y chromosome motifs in their genetic make-up.

Although the Olympic Committee has been 'sex-typing' women athletes since the 1968 Olympics, medical personnel

have until now used only a chromosomal procedure — known as the Barr-body test — that requires experience and sophisticated equipment for diagnosis. Beginning with the 1992 Winter Olympics in Albertville, France, however, the Olympic Committee wants to start sex-typing using the PCR technique, in which primers unique to the Y chromosome are used to detect the presence of the male-specific genetic material with better than 98 per cent accuracy. The procedure is easy enough to be done by a technician using a prepared kit of primers and common PCR equipment.

Although the Olympic sex tests are designed to screen out athletes with male musculature who are competing as females, a PCR test would also flag genetic variations that provide for no athletic benefit. In one variation, 'XY females' can have a defect in the Y chromosome which renders it inoperative, and therefore confers no athletic advantage over a normal XX female.

Estivill argues that the ease with which the PCR sex test can be conducted risks widespread sex testing in the absence of a clear idea of what the results actually mean. Although the Olympic Committee defines women as having two X chromosomes, there are several alternative chromosomal arrangements that still produce what society considers a woman — even if the scientific community is undecided.

This week the International Olympic Committee is expected to make a decision about what to do in Barcelona. Despite the rejection of such sex typing by many geneticists, the Committee is expected to have little trouble finding some laboratory to do the testing if it decides to continue with the trial.

Christopher Anderson

US RESEARCH POLICY

Ethics rule reconsidered

Washington

A US ADMINISTRATION official testified last week that he had been surprised by the storm of opposition prompted by a set of new proposed ethics regulations that would restrict the participation of government researchers in professional societies and that the government would redraft the proposal. Steven Potts, director of the Office of Government Ethics, told a congressional hearing on the subject that his office had received more than 1,000 responses to the proposed rules, more than he had seen on any other issue. Scientific societies had organized a letter-writing campaign to oppose the rules, which threatened to prohibit researches at the National Institutes of Health, the Department of

Energy and other government agencies from editing journals, organizing meetings or otherwise participating in society work on government time (see *Nature* 353, 5/5 September 1991). The societies argued that such participation is an important part of being a working scientist and does not compromise government integrity.

Potts said the regulations had been intended only to codify what his office believed was standard practice. But given the unexpected outcry, his staff was redrafting the rules, and would possibly release them for public comment again. This is expected to delay issuance of the final rules by about three months, to next spring.

Christopher Anderson

Plants, not pandas

RESEARCHERS working to preserve the genetic stock of rare and endangered plant species are meeting in Washington this week to plan a strategy for putting plant genetic resources on the agenda at the 1992 'Earth Summit', the United Nations Conference on Environment and Development to be held in Rio de Janeiro next June. 'Plant species on farmers' fields are being lost even faster than endangered animal species. They're not pandas, but they're vitally important', says Geoff Hawtin, director of the International Board for Plant Genetic Resources. He hopes to get a global plant preservation initiative from the Rio conference. Such an initiative would cost some \$300 million per year — between three and five times current spending on the preservation of agricultural genetic material.

C.A.

Not quite so dumb

US STUDENTS are nowhere near as bad as international comparisons of test scores suggest, a National Science Foundation (NSF) official writes in the new issue of the National Academy of Engineering journal *The Bridge*. Well-publicized declines in Standardized Achievement Test (SAT) scores, which are used for university admission, have resulted more from the fact that more students took the test and attended college than the effect of any US educational declines, writes Ira Rotenberg, an NSF program director for education. And the sobering results of a 1988 international test for 13-year-olds, which placed the United States last among six countries, are just as likely due to methodological errors stemming from the small and spotty sample size than real differences in education, Rotenberg argues.

C.A.

Power of the press

ONCE *The New York Times* notices new research, scientists are more likely to cite it, according to a new study by a University of California, San Diego researcher. Comparing citation counts of research reports published in the *New England Journal of Medicine* (NEJM), sociologist David Phillips found that work reported by *The Times* received 73 per cent more citations than other NEJM articles. To eliminate the possibility that this simply reflected the skill of the *Times* reporters in spotting good research, Phillips examined a three-month period in 1978 when much of the *Times* staff was on strike. The editors continued to produce a paper of record, including science coverage, but it was not sold to the public. The articles the *Times* covered during this period did not score significantly better than others in NEJM, proving, Phillips argues, that scientists are taking their lead from the press.

C.A.

US biotech in good health

Washington

The United States is still number one in the commercial exploitation of biotechnology, but progress in the development and marketing of new products made through biotechnology has been slower than expected, according to a report* released this week by the congressional Office of Technology Assessment (OTA). The report emphasizes that the key to maintaining a competitive position in the world marketplace depends less on the targeted sponsorship of biotechnology by the federal government and more on a strong research base.

In fiscal year 1990, the federal government spent more than \$3,400 million to support research and development in biotechnology-related areas, more than 80 per cent of which was awarded through the National Institutes of Health. According to OTA, biotechnology-based drug development is flourishing. To date, 15 drugs and other biological products have been approved by the US Food and Drug Administration (FDA), and more than 100 others are in various stages of development. And 1991 has been the best year since the stock market crash of October 1987 for biotechnology companies hoping to raise capital in the public markets.

Despite this rosy picture, OTA reports that the commercialization of biopharmaceutical products developed using biotechnology has been slower than expected and that the impact of biotechnology on other sectors of industry — agriculture, the chemical industry and the environmental business — has been limited.

OTA notes that six years after the Co-ordinated Framework for Regulation of Biotechnology was first proposed and four

years after it was made final, regulations for genetically modified pesticides and certain microorganisms are still not in place. This failure has led industry representatives to complain that the approval process as it stands is unclear and serves as a disincentive to investment.

Other policy issues singled out by OTA as relevant to US competitiveness include the need to structure coherent tax policies and adequately protect intellectual property rights. As markets in biotechnology become increasingly global, OTA says that protection of intellectual property is likely to become less of a domestic issue and more of an international one.

Despite the fact that in 1981 Japan's Ministry of International Trade and Industry singled out biotechnology as a key technology for the future, OTA reports that Japan's weak publicly funded research base has led companies there to look for research and training abroad, notably in the United States. OTA concludes that increased competitiveness cannot be assured by targeting alone.

OTA predicts that in the near term (5 to 10 years), the European Community, with its traditional strengths in pharmaceuticals and agriculture, poses the greatest threat to US competitiveness in biotechnology. However, a fragmented research effort, adverse public opinion surrounding biotechnology and the uncertain effects of recently introduced Community directives on the contained use and deliberate release of genetically modified organisms could hold Europe back.

Although some foreign companies are actively investing in US biotechnology companies, OTA says that approximately 75 per cent of all mergers and acquisitions

are between companies based in the United States, as illustrated by the \$660 million merger of Chiron and Cetus and, more recently, the acquisition of a 60 per cent stake in Genetics Institute by American Home Products. **Diane Gershon**

*US Congress, Office of Technology Assessment, *Biotechnology in a Global Economy*, OTA-BA-494 (Washington, D.C.: US Government Printing Office, October 1991).

HUMAN GENOME PROJECT

Tit for tat on patents?

London

IF US National Institutes of Health researcher Craig Venter is successful in his attempt to patent 337 partially sequenced human genes (see *Nature* 353, 485; 10 October 1991), international relations among genome researchers will never be the same.

Tony Vickers, who manages the UK Medical Research Council's (MRC) Human Genome Mapping Project, says that, should Venter's US patent application succeed, British researchers would be forced reluctantly to follow suit, ending the MRC's policy of making its cDNA sequences available to all researchers free of charge. A decision in favour of Venter's application would be "seriously prejudicial to the whole thrust of the international Human Genome Project," Vickers says.

"We don't believe that there's anything inherently patentable in one of these fragmentary cDNA pieces," says Vickers. But if the US Patent Office decides otherwise "the [British] government will say: 'What are you doing standing around?'"

The sequences produced by Venter are for fragments of human genes, each consisting of several hundred bases from a gene that may contain thousands. These partial sequences may yield enough information to guess at the function of the genes from which they come. Vickers believes the inventive step lies in discovering a gene's function, and putting this knowledge to use, rather than in simply producing a partial sequence.

A successful patent application by Venter would bring the ugly prospect of patent disputes dividing the international genome community. Under US law, patents are awarded to the first to invent, rather than simply the first to file an application (if the initial inventor files within one year of an other earlier application). Vickers is convinced that Venter must have sequenced some of the same genes contained in the MRC's sequence database.

In France, the home of Europe's other major cDNA sequencing effort, genome project officials are similarly taken aback by Venter's move. According to an aide to Jacques Hanoune of the French medical research agency INSERM, who runs the French government's genome project, the French are now considering applying for their own patents, although no decision has yet been made. **Peter Aldhous**

SOVIET SPACE

Mars projects still alive

Moscow

REPORTS of the demise of the Soviet space programme notwithstanding, the exploration of Mars remains a top priority. The plan is that there should be four consecutive visits to Mars by instrumented space probes in 1994, 1996, 1998 and 2001. The smaller of two Mars-landing vehicles, called Marsokhods, has already been tested successfully on the Kamchatka Peninsula, at an outstation of the St. Petersburg Transport Machine-Building Institute.

The volcanic terrain of Kamchatka is believed to resemble that of Mars, but the testing of the 75-kg vehicle was interrupted in the summer by torrential rain, of which there is none on Mars. Further tests are planned in the Mojave Desert in the United States in May next year.

The smaller vehicle is planned for the

1998 mission, and the larger, with a mass of between 400 and 500 kg, for the mission in 2001. Both are designed on the same principles, with six driven wheels mounted on a flexible chassis. The wheels, which carry terrain-matching devices on their surfaces for reliable traction, are conical-cylindrical in shape and so elongated that, on the smaller vehicle, there is no gap between them.

Vitali Vernigora, from the Babakin Research and Test Centre in Moscow and one of the designers, says the Mars landers are designed to navigate all terrains and to clamber over obstacles in their way. Because of the distance between the Earth and Mars, it will be necessary for them to function semi-autonomously on the surface of Mars, travelling along preset paths loaded into their computers. **Yuri Kanin**

France's peers overseas

SIR — Your amusing, but pessimistic, view (*Nature* 352, 176; 1991) of the outcome of our international inquiry into the reputations of the professors employed in the French university system is incomplete.

Your readers should know of the full scope of our attempt to establish a clear, acceptable and (we hope) accurate evaluation system for French university research and graduate studies. We have a scientific advisory council of 25 members, half of them from outside France, including several Nobel and Fields laureates, and chaired by Jean Marie Lehn. The council reviews ministry policy every six months, publishes its opinions and formulates recommendations on longer-term scientific policy.

Thirty groups, each of approximately 15 experts, review all proposals for contract money, advise on the right to award master's or doctoral degrees and evaluate candidates for the newly established *prime d'encadrement doctoral et de recherche*. (These are awards of an annual bonus of FF30,000 for four years to those who have been particularly successful in teaching and directing doctoral students and in producing and publishing innovative research.) The expert committees are similar in several respects to those established a few years ago to evaluate British science and to rank all British university departments. In addition, hundreds of French and foreign experts contribute evaluations of particular projects by mail.

Seven scientific directors and their staff use all this material to prepare the minister's decisions. The results are discussed with the universities. The principal outcome is a four-year contract between the ministry and each university, committing the state to a minimum level of funding. All evaluations and grants are made public.

The creation of this system has undoubtedly been eased because of the high priority given by the French Government to education and research. One obvious result is that the funds available to universities for research and doctoral studies have risen by more than 30 per cent over the past three years, reaching FF2,900 million in 1992. (The global budget, including the salaries of professors and other staff, is in excess of FF11,000 million.)

Our further attempt to provide experts with quantitative and external qualitative data must be judged against that background. We are now attempting to use citation indices and have also launched the worldwide inquiry campaign mentioned in your leading article. We do not intend to use the responses

directly in making decisions about the funding or promotion of individuals. We want to assemble the answers by department, university and discipline so as to form an opinion of the reputation that costly state-financed scientific research projects have earned abroad. We want to see ourselves as others see us.

Everybody knows that there is no perfect method of evaluating, even in broad terms, the quality of continuing research. Even if the many hundreds of responses to our letter include some along the lines that you predicted, the great majority have been independent straightforward opinions of the reputation of French science in particular fields, and as such have been useful and useable. We shall make a synthesis available as soon as possible.

I should like to take this opportunity to thank the hundreds of scientists outside France who have taken the time and trouble to provide us with serious and unprejudiced responses.

VINCENT COURTILOTT

*Ministère de l'Éducation Nationale,
de la Jeunesse et des Sports,
61–65 Rue Dutot,
75732 Paris Cedex 15, France*

Reductionism

SIR — Douglas B. Kell (*Nature* 350, 268; 1991) and Roland Dixon (*Nature* 351, 685; 1991) have commented on the ignorance of students concerning reductionism. Kell suggests that the subject needs investigation.

In 1980, in the Australian National University, 61 advanced or graduate students of zoology, all competent, answered a series of questions on the philosophy of biology. More than half held all biology to be ultimately reducible to physical science; but only seven were able to name this doctrine. Twelve of their teachers were invited to answer the questions; of the seven who did so, five accepted the principle of explanatory reduction but only one could recall its name¹.

The inability to name a concept is not in itself important. More notable is the ignorance, revealed by those students, of a number of issues of great practical significance. Their ignorance reflected their experience of university study. Most universities concentrate on producing graduates who know little outside the very narrow range of their special studies; and such a training encourages an unthinking reduction of complex phenomena to the concepts of a single subject such as biochemistry or genetics. Socially important problems can usually be solved only by applying the methods

and findings of several disciplines². When, therefore, such specialists try to apply their knowledge in practice, they meet questions that they are not equipped to answer.

The handicap imposed by specialism is especially evident when the attempt is made to apply the methods or findings of the natural sciences to questions concerning human societies^{3,4}. There is therefore much to be said for encouraging students and their teachers to study the uses and limitations of explanatory reduction; but attempts to introduce transdisciplinary university courses, in which this could be done, still arouse much hostility⁵.

S. A. BARNETT

*Australian National University,
Box 4, GPO, Canberra, ACT 2601,
Australia*

1. Barnett S.A., Brown V.A. & Coton H. *Studies in Higher Education* 8, 23–32 (1983).
2. Ashby E. in *The Sociology of Science* (ed Halmos, P.) (University of Keele, Staffordshire, 1972).
3. Kitcher P. *Vaulting Ambition* (MIT Press, Cambridge, Massachusetts, 1985).
4. Barnett S.A. *Biology and Freedom* (Cambridge University Press, 1988).
5. Barnett S.A. & Brown V.A. *Interdisciplinary Science Reviews* 8, 294–296 (1983).

Powdered milk

SIR — James Vickers (*Nature* 353, 103; 1991) is obviously not a reader of *The Times* (18 July), nor of the *Church Times* (26 July), or he would have seen that Nestlé has addressed the issues on which the Church of England General Synod reached its extraordinary decision. Moreover, unless he is a member of the Synod, he will not have seen the letter addressed to all its members by Mr Peter Blackburn, UK chairman and managing director of Nestlé, before it began its debate. Blackburn said that, in the interest of truth and fairness, the Synod would want to be in possession of all the facts, and went on: "We assure you that we have nothing to hide and would more than welcome an objective review of our record, including consultation with the World Health Organisation [WHO] and governments of countries in which our company operates".

As your readers will know, the Synod chose to ignore this offer. It is alarming to see members of the scientific community mouthing activist slogans as Vickers does. Where does he get the idea that WHO "has been condemning Nestlé's activities in this area for years"? Or that donating infant formula is designed "to create a market for the product among nursing mothers"?

May I suggest that instead of accepting activist propaganda at face value he does what any serious scientist should do, and finds out the facts.

F. X. PERROUD

*Nestlé SA, Avenue Nestlé 55,
CH–1800 Vevey, Switzerland*

Defence of diluted water

SIR — Is it an unease consequent upon its inept handling of the high dilution affair in 1988 that dictates that *Nature* cannot refrain from printing my name in close proximity to derogatory comments?

The latest example is H. Timmerman's letter (*Nature* 352, 751; 1991) which labels our work as "nonsense theories". He refers to my promise¹ "to [try to] publish in the months to come indisputable proof" for the high dilution effect and dismissively asserts: "I have not seen such a paper". It is symptomatic of the Dallasization of science that one scientist should dare publicly to imply that a colleague is a liar or incompetent or both without having had the decency to write to that person or adequately inform himself.

Even more surprising is the fact that *Nature* printed these comments knowing that it had refused an article reporting the production of our 1988 work under the direction of one of France's most respected biostatisticians (Professor Spira, INSERM U 292) and refrained from all comment when these data were recently published by the French Academy of Sciences² and, with similar ones, presented at the 1991 FASEB meeting^{3,4}.

A possible explanation for such odd behaviour on the part of a reputable journal is, ironically, to be found on the same page as Timmerman's letter. It is provided by E. J. Krowitz, who recalls Kuhn's concept of science as the triumph of expectations over reality.

In fact, our research has progressed. In more than 200 experiments on rat and guinea-pig isolated hearts, we obtained, using ten different highly dilute agonists, extremely significant vascular and mechanical effects. The cardiac effects were recently observed by an independent team of French pharmacologists. A note and an article are in preparation.

J. BENVENISTE

INSERM U 200,
32 Rue des Carnets,
92140 Clamart,
France

1. Benveniste, J. *The Lancet* II, 944 (1990).
2. Benveniste, J. et al. *C. R. Acad. Sci. Paris* 312 série II, 461–466 (1991).
3. Benveniste, J. et al. *FASEB J.* 5, A1008 (1991).
4. Hadji, L. et al. *FASEB J.* 5, A1583 (1991).

■ The article submitted to *Nature* to which the author refers was rejected on the advice of two referees, who argued that the statistical analysis was flawed. The publication of ref. 2 took the form of a reply to an earlier article in the *Comptes Rendus*. — Editor, *Nature*.

British science

SIR — Your Manifesto for British Science (*Nature* 353, 105; 1991) is incorrect to suggest that the Ministry of Agriculture's contribution to the research programme of the Agricultural and Food Research Council (AFRC) is zero. True, the ministry no longer funds near-market work, but it currently contributes some £35 million a year towards AFRC research programmes.

Your manifesto also suggests the transfer of academic support functions in the agricultural, medical and natural environment research councils to the Science and Engineering Research Council and the incorporation of their 'rumps' as agencies of operational departments. This would be an unfortunate development. The institutes include some of the most outstanding scientific centres in the country, for example the Medical Research Council Laboratory of Molecular Biology at Cambridge and the AFRC Institute of Plant Sciences that includes the John Innes at Norwich. Furthermore, the mixed economy of institutes and university-based units and research projects provides both larger facilities and multidisciplinary groups together with the more flexible and smaller project grants. The future health of our research depends on the close interactions between these different groups and scientific programmes that cross the boundaries. Our own policy at the AFRC is to bring institutes and universities closer together rather than to separate them.

The whole question of the reorganization of the research council system was considered in detail some two years ago and rejected by the government in favour of reconstituting the Advisory Board for the Research Councils. Its remit includes in particular closer collaboration across research council boundaries within the present structures. Nothing would be achieved by re-opening this debate.

TOM BLUNDELL
(Secretary)

Agricultural and Food Research Council,
Central Office, Polaris House,
North Star Avenue,
Swindon SN2 1UH, UK

SIR — It is the climate in British primary and secondary schools that has led to a dearth of aspirants to a scientific career and the inevitable consequences for the future of science in this country. Yet this comment passed tangentially at your public meeting on 13 September. The

Letters submitted for Correspondence should be typed, double-spaced, on one side of the paper only.

debate was tightly focused on the money, morale and means needed to undertake good science. Now that the "baleful influence" has left 10 Downing Street we are back to pleading with ministers.

It is disconcerting that, after all these years, the message from society is still misunderstood by a scientific community that should be listening, observing and drawing conclusions. The blame for the fact that society and its politicians are grossly ignorant of science cannot be effectively laid at their door. It is the responsibility of scientists to ensure that the case for research is well understood throughout the many strata of society, each in its different way. Research scientists have very special responsibilities and these are not limited to questioning the unknown in a systematic fashion. They extend to working to ensure that everyone is brought up to understand that science is the only way to safeguard the life, physical and cultural, that is so precious to us all. That has to be the true meaning of 'relevance'. Science has to be relevant to people.

The entire academic community has allowed an educational vacuum to be created all around itself and now research scientists, justifiably in need of public funding, plead for better understanding from those it has spurned in its obsessive search for excellence and commitment to short university courses. The programme of the Committee on the Public Understanding of Science may be conceptually correct but the audience, albeit raucous, is deaf. So, therefore, are ministers in a democracy. And so, it seems, is the scientific community.

D. A. A. FAGANDINI

6 Alieyn Park,
London SE21 8AE, UK

Direct action?

SIR — Neville J. Woolf (*Nature* 351, 10; 1991) proposes that the link between weak electromagnetic fields (EMFs) and cancer is indirect, through the formation of toxic chemicals in the air. But EMFs may well act on tissues directly by means of the common contaminant magnetite, ubiquitously present in the air as a ferromagnetic dust which, when inhaled or ingested, may be phagocytosed into cytoplasmic vesicles. When exposed to varying magnetic fields, phagocytosed magnetite dust may damage the vesicles and other structures and, being itself an auto-oxidation catalyst, may assault vital structures in the cell.

This model could easily be tested by exposing cells carrying magnetite dust in vesicles to varying magnetic fields.

STEFAN H. KON

PO Box 12391,
Albany, New York 12212, USA

Animals in research

SIR — The letter by David Wolpert (*Nature* 352, 466; 1991) concerning the “moral rights of an organism” being related to the presence or absence of human DNA is wide of the mark. Almost all living organisms live at the expense of other organisms. Certain species of trees grow more rapidly than others, thus winning the competition for sunlight. Chimpanzees in the wild have been seen to kill and eat smaller primates. Our homes stand on soil previously occupied by a wide variety of vertebrates and invertebrates. The farmer's plough kills or displaces innumerable living creatures. The use of animals in research has brought about the enormous advances of twentieth century medicine, both human and veterinary. Wolpert himself has benefited from experiments on animals, dating from his receipt of vaccines, developed in animals, against diseases that killed earlier generations. Wolpert's human DNA, and that of the rest of us, has given us our human brains so that we can conduct research humanely, for the benefit of both humans and animals.

Wolpert's statement that “the vital organ for vivisectionists is in fact the wallet” is truly offensive. There are active groups of patients and their families (Incurably Ill for Animal Research; Canadians for Health Research) who

know first hand the benefits of biomedical research, and who also recognize that much more knowledge is required to conquer diseases at present incurable. Most of our colleagues in the field of biomedical research are concerned with the prevention and cure of disease, many of them entering the field because of first-hand knowledge of the effects of disease and injury on family and friends. Wolpert might ask himself whether his research at the Los Alamos National Laboratory has benefited his fellow human beings, or animals, anywhere near as much as has biological research.

H.J. RALSTON III
DIANE DALY RALSTON

*Department of Anatomy,
School of Medicine,
University of California, San Francisco,
California 14143-0452, USA*

SIR — David Wolpert asserts that “the vital organ for vivisectionists is the wallet”, and that the scientists most “incensed by the animal rights movement are those whose careers it threatens”.

His assertion about money is ‘below the belt’. We could be making much more money looking down the throats of children and pronouncing them ill. Instead, some of us have chosen to extend the frontiers of biology, even when it includes the messiness of using

large fraction (29.9 per cent) said that they felt there were circumstances in which a physician would be justified in causing a patient's death.

Religious beliefs and a concern for the legal consequences of their actions appeared as the factors that had the greatest influence on physicians' decisions in treating dying patients.

Although limited resources prevented us from meeting some of the more rigorous standards of social research, our findings can be said accurately to reflect the actions and attitudes of the cohort we sampled. They suggest that the extent of physician involvement with dying patients should be more comprehensively explored and the factors influencing near-death clinical decisions thoroughly defined. Without such data, objective discussion will be buried by emotional argument; and the publication of reliable data showing the true extent of physician-assisted death may be needed in order to force the issue onto the public agenda. Physicians appear reluctant to break the ‘conspiracy of silence’ on the issue.

MAC OVERMYER

*Physician's Management,
7500 Old Oak Boulevard,
Cleveland, Ohio 44130, USA*

nonhuman animals in our work.

To the second half of his argument I must agree. We who work with animals find the animal rights organizations irritating, illogical and sometimes dangerous. When they demand an end to experimentation with animals, it strikes the same nerves in me as in a physicist who is told that nuclear reactors should be banned or enquiries into the nature of matter itself should be halted because these things are ‘intrinsically unsafe’ or ‘go against the will of God’. If we don't defend our research, who will (or should)?

Most of Wolpert's letter concerns the argument that, “for vivisectionists, the primary determinate of moral rights of an organism appears to be whether or not the organism has human DNA”. I certainly hope not. Last month, I killed millions of cells containing human DNA. It is not the DNA that is important, but the effect of it. Humans have (a few) genes, different from those of any other animal, giving our species the ability to ‘change the world’. When rodents build something like the Parthenon, write a poem, or make fire, I shall be happy to give them the rights reserved for members of our own species.

Those of us who must use animals in our research must work with what evolution has provided. Theories, mathematical models and tissue cultures are a good start to understanding the cause and (possible) cure(s) for disease and dysfunction. However, at some point an animal is required to address the question “Does this work?” The animal rights fanatics would prefer that the research animal be *Homo sapiens*. Ultimately, it is. I believe no one is braver than the fellow who rolls up his sleeve for that first injection, swallows the first pill, or undergoes the first operation involving a novel method. I also believe no one is crazier than one who would submit to those experiments without first knowing the effects on another mammalian species.

Finally, I would like to explain that I have been fortunate not to have had to resort to the use of apes in my research. Like most animal researchers, I work with nonprimates. I have observed some chimpanzees whose social behaviour (personality?) could be described as ‘human’. I have also met humans whose brains are so dysfunctional that they are sometimes described as ‘vegetables’. The thought of chimps behind bars with electrodes in their heads disturbs me. So does the thought of those unfortunate humans. What is one to do?

JAMIE LOVE

*Institute of Animal Physiology
and Genetics Research,
Edinburgh Research Station,
Roslin, Midlothian EH25 9PS, UK*

Final exit

SIR — In the leading article “Final Exit” (*Nature* 352, 553; 1991), you advanced the position that “the ‘ethic’ of using technology to prolong life when life is ebbing should give way to the ethic of individual choice, including the right of physicians to choose to assist the dying”.

Many US physicians have already adopted this ethic and physician-assisted death is one of society's better-kept secrets. At the beginning of 1991, this journal (*Physician's Management*) conducted a survey of 2,000 primary-care physicians to find out their response to ‘end-of-life dilemmas’. We also sought to identify factors that influenced clinical decisions in such situations.

Analysing the data gathered from nearly 500 respondents, we found that nearly all (92.4 per cent) had issued “do not resuscitate” orders for dying patients and that more than half (58.5 per cent) had ordered the removal of life-sustaining technology. More surprising, some 9.4 per cent responded that they had “deliberately taken clinical actions that would directly cause a patient's death” and 44.8 per cent said that they had deliberately taken clinical actions that would “indirectly” cause death. A

A thesis for all seasons

Lars H. Breimer and Dimitri P. Mikhailidis

Much could be gained by a reform of the system by which doctoral degrees are awarded, not least because of the single European market, due in 1993.

A DOCTORAL degree is the first stage of the journey towards a career as an independent researcher. Candidates need to be guided not only towards thinking independently but also to developing the communication and organization skills necessary for running their own research groups. Much of today's biomedical research is done in the United States and the United Kingdom, yet the nature of the doctoral system in these countries means that the results of research projects are not necessarily widely available and that candidates spend an inordinate amount of time on the mechanics of thesis production. Therefore, we suggest that candidates are allowed to submit theses based on published work and that theses are published in an international thesis journal. Furthermore, the lack of a unified procedure and accreditation of doctoral work could hinder free movement of people within Europe because of local attitudes to foreign degrees and titles. This problem is particularly important, of course, for those entering or returning to industry or to a profession such as medicine or veterinary science. Reform of the procedures for awarding research-based degrees must therefore address on the one hand the problems of thesis examination and, on the other, training and use of the time allowed for research.

Current rules

At first glance, the systems for training in research and awarding doctoral degrees differ considerably between countries, especially in the nature of the thesis proper, its examination, requirements for course work, and in the general availability of the thesis. For practical purposes, however, essentially two procedures for awarding doctoral degrees are used throughout the world. In a few countries, illustrated here by Sweden, a thesis is normally based on published papers, but in most, including the United States, United Kingdom (and its former empire) and several other European countries, a thesis normally does not contain published papers.

US and UK theses are laboriously put together by the candidates themselves and published in small numbers (typically four copies). The candidates may have published papers but these are not integrated in the thesis; methodology and results are therefore included in detail.

The rules of eligibility also vary between countries. In the United States, a student is generally accepted into a graduate school and starts by undertaking two years of course-work, including practical rotations between groups, for a master's degree before proceeding to a PhD project with one of the groups. Thus, it usually takes at least five years to achieve a doctorate. In most other countries, the candidate starts on the practical work immediately. In the United Kingdom, there is no formal requirement for course-work but candidates as a rule register initially for a master's degree which may never actually be written.

By contrast, a Swedish thesis consists of four or five papers that have been published in international refereed journals, combined with a substantial review and discussion of up to 10,000 words. There is no need to repeat descriptions of technical and practical details or of results. The thesis is professionally printed (about 150 copies is typical) and distributed at the faculty's expense. To be entitled to submit a thesis the candidate must as a rule complete three, usually practical, courses to broaden her or his scientific base. These 'doctoral courses' are often residential, held at other universities and can result in potentially useful informal contacts.

In Japan, a PhD can be acquired by either route. Briefly, the candidate can enter a programme that is not dissimilar from that in the United States in that after the basic university degree the candidate undertakes a 2-year master's programme and proceeds to 3-4 years of PhD studies. Alternatively, a thesis of

about 20,000-25,000 words may be based on 4 or 5 papers published in international journals.

In some countries, including Portugal, the undergraduate courses often take longer and the system also puts great store by the master's degree. In consequence, PhD students are usually much older and may already be established in their career before starting their doctoral work.

In a limited survey, we inspected the successful PhD theses submitted at our medical school during two decades. We found that in those theses containing bound published works, the average number of papers included steadily increased from 2.1 per thesis during 1971-75 to 3.6 per thesis during 1986-90 (see table). This latter number is comparable to the 'expected' average for a Swedish or Japanese thesis submitted on the basis of published works.

In many European countries and in Japan, theses are generally publicly defended before the examiners, the audience being allowed to ask questions. In the United Kingdom and United States, however, the oral examination is held *in camera*.

A superleague

Thus, it seems that the organization of North American baseball, where two main leagues play a game under different rules yet meet in a final (quaintly named the World Series) is a better analogy for the state of doctoral degrees than the organization of English soccer, where the Football Association and League administer different parts of the game but the rules are the same. Unified

SUCCESSFUL PhD THESES AT THE ROYAL FREE HOSPITAL SCHOOL OF MEDICINE

Years	Total number of theses	Theses with papers	Total number of papers	Ratio of theses with papers to total theses	Number of papers per theses with papers
1971-75	27	8	17	0.3	2.1
1976-80	17	6	15	0.35	2.5
1981-85	20	8	28	0.4	3.5
1986-90	39	13	47	0.33	3.6

We inspected the successful PhD theses at the Royal Free Hospital School of Medicine in 1971-90. The total number of theses which included published papers was noted as was the number of papers included in each of these theses. The ratio of theses with papers to the total number of theses and the number of publications per theses which had included papers were calculated. The table shows that although the proportion of theses which include papers has not changed over two decades, the number of papers included in each of these theses has almost doubled.

regulations allow a ready comparison of the quality of theses between countries and will facilitate the free movement of graduates between countries.

We propose that a new system of preparing and examining a doctoral thesis should be introduced in the United Kingdom on a trial basis. This system would operate in parallel with the existing one, being a prototype for a unified European doctorate. It would address the questions of impartiality, improved use of time, supervisor commitment, increased availability of the thesis, and cost. Briefly, we feel that these goals can best be achieved through a new type of thesis; a public oral examination; and publication of (at least part of) the thesis.

A new type of thesis. The thesis should consist of a compilation of at least four 'first-author' papers published in international, refereed journals recognized to be relevant to the field, together with a 5,000–10,000-word review of the subject (the 'thesis overview'). A list of 'approved' journals could be published. The Swedish and Japanese systems and our limited survey suggest that a four-paper requirement is reasonable. The thesis overview should explain the background and reasons for undertaking the work, and also how the results contribute to the field.

One advantage of the inclusion of published papers is that the candidate's work has already been peer-reviewed. Any correspondence in the original journals relating to the papers forming part of the thesis would be included, with comments. Another advantage is that the candidate does not need to spend an excessive time on the mechanics of thesis preparation: indeed, thesis 'padding' should be severely criticized and brevity lauded.

Public oral examination. The doctorate must be publicly defended. This favours impartiality, and would ensure correct behaviour between examiners, candidate and supervisor. The audience should also be entitled to question the candidate. The public examination would render unnecessary the need for alternative methods of ensuring impartiality, for example by increasing the number of examiners so as to reduce the risk of biased selection. Occasionally, however, where for example the work straddles several fields or involves special techniques, additional examiners may be needed.

Thesis publication. To allow audience participation, the thesis must be freely available to the scientific community. Thus, it must be published as a supplement to an appropriate journal, as a review article or as a monograph. The cost of publication and reasonable distribution should be defrayed by the

candidate's department.

We suggest the creation of a thesis journal in which the overview of each candidate's thesis would be published not less than a month before the public examination. The journal would be carried by all libraries in the United Kingdom and be offered (free or at low subscription rates) to libraries worldwide. Given that theses are usually excellent, up to date and unbiased reviews of a field, this could be seen as a low-cost service to libraries in developing countries, which often can afford to subscribe to only a few journals. A wide circulation would also make the journal attractive to advertisers, which would help offset the cost.

The thesis journal would allow specialists to decide whether to attend the oral defence of a particular thesis, or whether to communicate in writing with the examiners so that their questions could be asked. The outcome of theses examinations would appear in subsequent issues of the journal. Prepublication of the final title and synopsis of a thesis (say about six months before submission) might improve research efficiency by reducing overlap with new projects. Of course the journal, or at least the 'pre-submission' section, could be in an electronic rather than printed mode.

Comment

We believe our system ensures objectivity and fairness. Work would have been subjected to refereeing before publication of the papers, to discussion by the scientific community, and to public examination. The thesis journal ensures that the candidate's effort is not wasted but that what is probably a most up-to-date and objective review of the literature is readily available.

Our proposed system will encourage candidates to publish their work. The ability to write papers concisely and clearly, together with 'hands-on' familiarity with the publication process, should form an essential part of postgraduate training. Moreover, candidates will be encouraged to publish a few papers in quality journals rather than many papers in less good journals.

The main objections to our scheme are practical and financial. The examination would have to be held in a lecture theatre and might well take longer than at present. Publishing a thesis may be considered an unnecessary expense. But there is no reason why an oral examination should take significantly longer if held in public as opposed to *in camera*, provided that the chairperson is efficient. It is unlikely that there would be many questions from the audience (after all, the examiners are experts in the field) and the proceedings could be concluded within a practical time-limit. The

public nature will ensure a fair hearing and a reasonable standard, with less risk of candidates being let through because of examiner deference or failed for 'political' or other unfair reasons.

Shifting of the financial burden of producing the thesis from the candidate to the department will help to ensure a more careful selection and supervision of doctoral students. Admittedly, it would be unreasonable for the department to have to defray a second edition of the thesis should it require revision.

The art of public speaking, another component of research training, will have to be developed if the candidate is to defend his work in front of an audience. The present UK system allows a candidate who has never published or spoken publicly to succeed. A crude analogy of this latter event would be to allow a physician to practice although he has never actually examined a patient. The proposed system will produce a new breed of graduates who are more skilled in communication and management as a result of exposure to these issues.

As in any training system, there is a need to safeguard against discriminating in favour of large research groups, in which young postdoctoral students are hardly trained to think originally or work independently. Indeed, the current UK/US system does allow a small research team to train a doctoral student successfully and also gives the supervisor the option of letting a candidate undertake a speculative project. Our published-paper system could favour 'safe' projects and speculative work might be ignored. One safeguard would be that the new system must coexist with current procedures so that an option for more speculative projects and negative findings remains. However, the new system will not help candidates whose work does not contain methodological or conceptual developments, as such work is unlikely to be published.

Nevertheless, we believe our proposals will be more rewarding to the doctoral candidate, not least because his or her work will become known to a wider audience than at present. The reform should also result in better use of research time, improved training, fairer examination and unified regulations. This should simplify comparison between participating countries and facilitate movement of scientists, particularly important in the European Community beyond 1993. □

Lars H. Breimer and Dimitri P. Mikhailidis are in the Department of Chemical Pathology and Human Metabolism, Royal Free Hospital School of Medicine, University of London, Rowland Hill Street, London NW3 2PF, UK.

We thank Dr K. Kodama and Professors A. F. Winder and A. J. Zuckerman for advice and encouragement.

The semantics of plane-mirror inversion

Those who ask why mirrors "invert right and left", but not "top" and "bottom", must not look for help in physics textbooks, but in dictionaries that define the use of words and in an understanding of reality.

It is, of course, a commonplace that people will say that an image in a plane mirror is inverted in the sense that 'right' is replaced by 'left'. You only have to look, and you will see! Hold up your right hand, palm forward, in front of a plane mirror, and you will see what looks like a left hand facing back at it. Moreover, the image is clearly not just a flat two-dimensional cut-out; if you keep the hand fixed, but move your head, the mirror will reflect light from a slightly different subsurface of your hand, and so the image will seem itself to be three-dimensional.

If all this is accepted, then an uncomfortable question usually arises: if reflection in a plane mirror inverts right and left, why does it not also invert top and bottom? Often one comes across explanations based on the assertion that the horizontal plane is distinguished from the vertical in that human beings have two eyes which, between them, define a unique horizontal plane, but that there is no means by which a unique vertical plane is defined by the physiology of the onlooker. There are all kinds of philosophical objections to arguments of that kind, but the most telling way of turning them is the simple observation that even a one-eyed man will see the reflection of his right hand in a plane mirror as what seems to be a left hand. Another is to suppose that the observer reclines horizontally in front of a plane mirror; again there is no inversion of top and bottom, but again a right hand is reflected in an image looking like a left hand.

So what is the explanation? E. C. Horsfield from the University of Durban has now persuaded the *European Journal of Physics* to let him publish a solemn discussion of the issue (12, 207; 1991) which appropriately begins with quotations from well-known undergraduate texts asserting that "images in a plane mirror have right and left interchanged". They are all, of course, "fallacious".

Horsfield's view is that the underlying error in these statements is that their authors confuse "physics and psychology" — the reflection of light in mirrors and what observers construe of the images they see. And, correctly, he points out that the only possible effect of a plane mirror that may be taken as inversion of any kind is that in which an array of points on a straight line perpendicular to the surface of the mirror will appear, in reflection, to be inverted. There is no lateral displacement

of any kind, such as would be needed to bring about left-right or even top-bottom inversion, but only inversion in a direction perpendicular to the plane.

That point is easily verified. If $\mathbf{X}$ is a vector in front of and anchored into the surface of the mirror, and $\mathbf{n}$ a unit vector perpendicular to the mirror anchored at the same point, the component of $\mathbf{X}$ perpendicular to the mirror is $(\mathbf{X} \cdot \mathbf{n})\mathbf{n}$ and the component in the surface of the mirror is simply $\mathbf{X} - (\mathbf{X} \cdot \mathbf{n})\mathbf{n}$. The effect of reflection is simply to replace $\mathbf{n}$ with $-\mathbf{n}$, so that the reflection of the vector $\mathbf{X}$ is simply $\mathbf{X} - 2(\mathbf{X} \cdot \mathbf{n})\mathbf{n}$. All that means is that the image of a point a certain distance in front of the mirror is displaced by twice that distance towards the other side of the surface, and that there are no lateral displacements.

How, then, does the illusion of left-right displacement arise? Horsfield goes on to argue that, whatever the mechanism, the illusion is incomplete. A person with hair parted on the left who looks into a mirror will see an image of what appears to be a person with hair parted on the left-hand side. But if the image is to be taken as that of the face of a real person, the parting of the hair will be on that person's right. So the image does not correspond exactly to the appearance of the person in front of the mirror. And the lateral inversion occurs only in "the observer's mind".

Horsfield attributes the root of the trouble to people's common determination to personify the images they see as reflections in mirrors. By way of antidote, it could be worthwhile for all of us to compare our reflections in a mirror with a photograph taken by a camera, when the two images would again suffer left-right inversion. But the only long-term remedy is to believe the truth — that all images are virtual images, without reality, just as are the virtual images supposed in the elementary optics textbooks to be formed in front of (not behind) convex spherical mirrors.

Astronomers and those who read their journals are well used to the difficulty: photographs of the sky habitually have 'East' on the left and 'West' on the right, yet nobody seems to turn a hair. The verbal explanation is simple. The mirror image of the sky is simply the projection onto a mirror of what seems to be a surface full of stellar objects. One's view of the image is then what one would obtain if it were possible to position oneself behind the

stellar surface to observe the arrangement of the objects. It is as if the object in view is flattened onto the mirror, and one is enabled to look at it from behind. This is also what happens with the image of one's face in a plane mirror; the mirror captures a front view of one's face and also enables one to look at it from behind. The trick amounts to nothing more than the inversion of the perpendicular components of all vectors in the mirror surface.

Grand philosophical questions naturally arise, but should be dismissed. There is, for example, the issue of 'reality', much discussed by people such as Hume and Kant. Is the image in a mirror real? Clearly it is not. The most that can be said is that an observer can reconstruct the supposedly real object that generates the image once the rules are known — that the left seems to appear on the right, for example.

But even that does not put an end to philosophical objections. Who, after all, can know what exactly happens within our heads in the direct perception by the eyes of objects placed in front of them? However perception works, may it not be likened to a system of mirrors that leads, willy-nilly, to the formation of an image by means that are in principle no different from the formation of an 'inverted' image in a plane mirror? And while all the members of a group standing round, say, a table will be tempted to vote for the proposition that "That is a table", there is no objective way of demonstrating that all of them are seeing the same thing.

Historically, arguments like these prompted Ernst Mach's positivist view of reality — that the only attributes of a physical object with significance are its measurable attributes. But even that is only half the battle. Have not generations of young people been drilled at school to measure the 'position' of a mirror image by sticking pins in an optical bench, looking for the position at which the image appears not to move? The truth is that these experiments do not measure an attribute of the image, but rather identify a fixed point from which all rays reflected in a mirror from a common origin appear to diverge. The trouble is simply semantic — that of pretending that the image is an object. The remedy is not the obvious abolition of 'image', for the word is altogether too useful for that. The best course is simply to remember that all images are unreal.

John Maddox

Peptides of fortune

Charlie Janeway

In the television show *Wheel of Fortune*, participants try to guess a phrase based on a few letters placed at specific positions in words of defined length. Precisely the same principle has now been applied by Bevan and colleagues on page 852 of this issue¹ to identify a specific peptide that allows protective cytotoxic (killer) T cells to identify and destroy cells infected with the intracellular bacterium *Listeria monocytogenes*. This finding and others open the way to the rapid identification of peptides that might be used to create synthetic vaccines; such vaccines would be invaluable in the fight against the many infectious diseases that continue to exact a heavy toll on human populations².

Immunological memory

Vaccination works by inducing a state of lasting immunity known as immunological memory. Memory requires the activation of T lymphocytes specific for the infectious agent. The T cells detect infection within cells by recognizing peptide fragments derived from the pathogen. These fragments are transported to the cell surface bound in a deep groove on the outer aspect of a molecule encoded in the major histocompatibility complex (MHC)^{3,4}. CD4⁺ T cells, which activate B cells and macrophages, recognize MHC class II molecules binding peptides of about 15 amino acids derived from proteins degraded in cellular vesicles. CD8⁺ T cells, which kill cells infected with viruses or intracellular bacteria such as *L. monocytogenes*, recognize MHC class I molecules binding peptides of eight or nine amino acids derived from proteins degraded by cytosolic proteasomes and transported into the endoplasmic reticulum by specialized transporter molecules of the ATP-binding cassette family⁵. Individual amino-acid side-chains, located at precise positions along the peptide, bind in deep pockets that line the peptide-binding groove of MHC class I molecules⁶. The position of these pockets and the amino acids that line them are different for each MHC class I allelic variant; thus, each MHC class I molecule binds a different set of peptides. This MHC polymorphism causes differences between individuals in resistance to specific infections⁶.

The stable binding of peptides to MHC class I molecules allows their direct characterization, and Van Bleek and Nathanson⁷ recently sequenced a single peptide of the influenza virus eluted from a specific MHC class I molecule isolated from infected cells.

The peptide was recognized by CD8⁺ killer T cells in mice of that genotype, demonstrating that peptides identified in this way are functionally relevant. This approach extended earlier studies of Rammensee and colleagues⁸, who from whole-cell lysates had isolated small peptides that could sensitize cells for killing by specific T cells. Rammensee's group subsequently showed that pooled peptides of eight or nine amino acids, eluted from a purified MHC class I molecule, had dominant amino-acid residues at one or more positions⁹; peptides from different MHC class I molecules yielded distinct amino-acid motifs. This result has been confirmed in a specific case by Jardetzky *et al.*¹⁰ by sequencing individual peptides isolated from the MHC class I molecule HLA-B27.

Bevan's group¹ used the sequence motif X-Y-X-X-X-X-X-X-(V,I,A,L) (single-letter code) for H-2K^d-associated peptides to identify a specific peptide of nine amino acids derived from *L. monocytogenes*. This peptide is derived from listeriolysin, the bacterial protein that enables *L. monocytogenes* to enter the cytoplasm. The CD8⁺ killer T cells were known to recognize a fragment of listeriolysin bonded to the mouse class I molecule H-2K^d. Having produced all 11 nonamers with this motif, they found that one efficiently sensitized H-2K^d target cells for lysis by their cloned T-cell line. A similar finding has also been reported by Rammensee and colleagues¹¹ who identified a nonamer from a synthetic peptide of 19 amino acids that had previously been shown to confer sensitivity on target cells. So knowing the peptide motif of a specific MHC class I molecule can allow rapid identification of specific peptide epitopes from antigenic proteins of known sequence.

How great is the simplification that this procedure allows? It is theoretically possible to make 5×10^{11} peptides of nine amino acids. Jardetzky *et al.*¹⁰, on the basis of 14 peptide sequences known to bind HLA-B27, can define a motif for that molecule that could accommodate over 1.3×10^7 peptides, indicative of the remarkable capability of any MHC molecule to present a wide range of distinct structures. This motif reduces the complexity of possible peptides by a factor of about 40,000. Nevertheless, the identification of one specific peptide is still a daunting task unless one has a fairly small target, such as the 19-amino-acid peptide studied by Rammensee and colleagues¹¹ or the challenging 59K product of the listeriolysin gene¹. So

although it is a marked improvement on random searching for peptides, the *Wheel of Fortune* approach still involves considerable work.

An alternative approach to identifying critical antigenic peptide epitopes, one which generalizes that employed by Van Bleek and Nathanson⁷, is suggested by the work of Jardetzky *et al.*¹⁰ for MHC class I and of my own group¹² for class II. The method involves purification of MHC molecules, isolation of the bound peptides, and separation of the peptides by high-performance liquid chromatography. Surprisingly, direct sequence analysis of several of the peaks identified single peptides which were derived from proteins abundant in the appropriate cellular compartment. Some of these were derived from an infectious agent, and the identified peptides were immunogenic when tested *in vivo*. This approach allows the identification of any abundant peptide, does not require the use of radiolabelled amino acids and, most important, does not require knowledge of either immune reactivity or the genome of the infectious agent. But it remains to be seen how well it will work with unknown infectious agents.

Clinical application

These new techniques, which were originally developed to answer basic questions about the nature of antigen presentation, are now ready for clinical application. They should speed the search not only for peptides for use in vaccines, but also for autoantigens¹³ important in clinical medicine. A combination of direct peptide sequencing and the use of specific motifs in known antigenic proteins should generate much new information about the involvement of such peptides in specific infectious and autoimmune diseases. □

Charlie Janeway is in the Section of Immunobiology and Howard Hughes Medical Institute, Yale University School of Medicine, New Haven, Connecticut 06510, USA.

1. Pamer, E. G., Harty, J. T. & Bevan, M. J. *Nature* **353**, 852-855 (1991).
2. Bloom, B. R. *Nature* **342**, 115-120 (1989).
3. Robertson, M. *Nature* **353**, 300-301 (1991).
4. Hackett, C. J. *Nature* **349**, 655-656 (1991).
5. Madden, D. R., Gorga, J. C., Strominger, J. L. & Wiley, D. C. *Nature* **353**, 321-325 (1991).
6. Hill, A. V. S., Allsopp, C. E. M., Kwiatkowski, D., Anstey, N. M., Twumasi, P., Rowe, P. A., Bennett, S., Brewster, D., McMichael, A. J. & Greenwood, B. M. *Nature* **352**, 595-600 (1991).
7. Van Bleek, G. M. & Nathanson, S. G. *Nature* **348**, 213-216 (1991).
8. Rötzschke, O., Falk, K., Wallny, H.-J., Faath, S. & Rammensee, H.-G. *Science* **249**, 283-287 (1990).
9. Falk, K., Rötzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290-296 (1991).
10. Jardetzky, T. S., Lane, W. S., Robinson, R. A., Madden, D. R. & Wiley, D. C. *Nature* **353**, 326-329 (1991).
11. Rötzschke, O. *et al.* *Eur. J. Immunol.* (in the press).
12. Rudensky, A., Preston-Hurlburt, P., Hong, S.-C., Barlow, A. & Janeway, C. A., Jr *Nature* **353**, 622-627 (1991).
13. Ong, B. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **88**, 7343-7347 (1991).

Why did Halley hiccup?

Paul Weissman

DOES anyone remember Halley's comet? It was in all the papers about five years ago, but since then not much has been heard about it. That is, not until the comet flared into brightness again in February of this year¹. The only problem was that the comet was over 14 astronomical units (14 AU, about 2,100 million km) from the Sun, between the orbits of Saturn and Uranus, a region of the Solar System so cold that it was believed that a comet nucleus would be totally inert there.

Astronomers at the European Southern Observatory had been dutifully following Halley as it moved outward from the Sun, heading towards aphelion (furthest from the Sun) beyond the orbit of Neptune in the year 2023. The comet had faded to 25th magnitude, about 30 million times fainter than the faintest stars visible with the naked eye, and showed no signs of activity. But on 12 February, Halley was discovered to have brightened by a factor of about 300, and a dust coma was observed around the nucleus, extending more than 100,000 km from it (see figure). Estimates of the rate of coma expansion showed that the outburst must have occurred in late January and activity persisted well into March.

Specialists are at a loss to explain this strange behaviour. One simple (though rather improbable) suggestion² is that a meteoroid or small asteroid, between 2.6 and 57 metres in diameter, collided with Halley's nucleus. This impact would not have destroyed Halley; rather it would simply have created a small crater on the surface of the nucleus, blasting off a cloud of ice and dust. But not much is known about the flux of small bodies among the outer planets, and only a few large asteroids and comets have been identified in that region. Unlike the asteroid belt between Mars and Jupiter, there are no dynamically stable zones between the orbits of the outer planets where material could be stored over the age of the Solar System³. Some cometary debris might be temporarily resident in this region, having been scattered inward from the hypothesized Kuiper belt, a remnant population of icy planetesimals that is thought to exist beyond the orbit of Neptune⁴, and which may be the source of the short-period comets.

Halley's relatively small cross-section (8×16 km) does not make it a very substantial astronomical target: a small comet or asteroid orbiting at 14.3 AU would have a probability of about 2×10^{-18} of striking the comet as it passed

by. Also, the high inclination of Halley's orbit argues against an impact, as most material in the Solar System (except for comets) tends to lie in orbits close to the ecliptic plane. Halley was 4.3 AU below the ecliptic plane at the time the outburst was discovered.

A somewhat more complex (though still unlikely) mechanism, from Intrilligator and Dryer⁵, is that shocks from massive solar flares may have impacted the comet and somehow initiated a breakup of the nucleus's nonvolatile

IMAGE
UNAVAILABLE
FOR COPYRIGHT
REASONS

Halley in outburst, February 1991.

surface crust, exposing volatile ices below. A series of solar flares in December and January, just before the outburst, were among the largest ever observed.

A serious problem with this proposal is the low energy density in solar-flare shocks, which is less than the energy density in sunlight. Intrilligator and Dryer calculate that the shock pressure impacting the nucleus surface at 14.3 AU would be 3×10^{-9} dynes cm^{-2} , hardly a comet-shaking force. The physics of cometary surface crusts is not well understood, but it is difficult to believe that a crust layer that was roasted during Halley's closest approach to the Sun, at only 0.587 AU (inside the orbit of Venus), would react significantly to such a small shock. And why didn't the comet react in the same way during a series of similarly large solar flares in late 1989, when Halley was closer to the Sun?

Perhaps the most likely explanation of Halley's strange behaviour comes from a talk by J. Klinger (Laboratoire de Glaciologie et de Géophysique de l'Environnement) at a meeting* earlier this summer. Klinger combined two older ideas: first, that the outburst was from a pocket of carbon monoxide ice, and second, that the outburst was triggered by the transition of amorphous water ice to crystalline ice, many metres below the nucleus

surface. Carbon monoxide was the second most abundant volatile observed in comet Halley (after water)⁶ and is one of a few volatiles that could still be active at the 70-K temperatures expected on the comet nucleus 14 AU from the Sun.

If comets were formed cold, as current theory predicts, then the water ice was probably formed in an amorphous state. Amorphous ice puts out heat as it turns into crystalline ice at a temperature of about 140 K. Computer simulations⁷ of heat transport in cometary nuclei show that this transition first occurs at about 5 AU on the inbound journey, as the comet first approaches the Sun from the distant Oort cloud, transforming several to tens of metres of surface ice to the crystalline state. On later orbits, the transition occurs well after perihelion, at about 8 AU outbound, as the thermal heat wave from perihelion slowly penetrates into the buried amorphous-crystalline interface. Eventually, all the ice in the nucleus is converted. But, because Halley's eccentric orbit keeps it well away from the Sun for most of the time, Halley is one of the few short-period comets which is expected still to have amorphous ice in its cold interior⁸.

Klinger's explanation is still far from complete in explaining why Halley's outburst occurred so far from the Sun and in describing just how carbon-monoxide-rich pockets of ice and/or gas might form. But its strength lies in the fact that it uses cometary processes that have already been studied and understood somewhat, and that it does not invoke events of exceedingly low probability.

Comet Halley is not alone in its strange behaviour at large solar distances. In 1988, astronomers noted a brightening of asteroid 2060 Chiron, and the following year detected a visible coma around the object^{9,10}. Chiron is one of a very few known objects in the outer Solar System whose orbits cross those of planets; its motion takes it across the orbit of Saturn and almost out to Uranus, but no closer to the Sun than 8.4 AU. A search for Chiron on pre-discovery plates by S. J. Bus (Lowell Observatory) and colleagues showed that it was even brighter near aphelion in 1970. As with Halley, carbon monoxide has been suggested as the likely volatile driving the activity on Chiron. Recent dynamical simulations by Hahn and Bailey¹¹ showed that Chiron, estimated to be 160–340 km in diameter, has a reasonable probability of being perturbed into an Earth-crossing, short-period cometary orbit in the next 10^5 years, or that it may have already been in one in the past.

A bit closer to the Sun, the unusual short-period comet Schwassmann–

European Southern Observatory

* Asteroids, Comets, Meteors, 1991, Flagstaff, Arizona, 24–28 June 1991.

Wachmann 1 is in a low eccentricity orbit between Jupiter and Saturn. The comet nucleus, estimated to be 30–40 km in diameter, shows frequent outbursts and near continuous, low-level activity¹². This greater proximity to the Sun allows for carbon dioxide to be the driving volatile, or even perhaps water if the nucleus's rotation pole is tipped over, facing the Sun. It has been suggested that much of the unusual activity in comets can be explained by 'seasons' resulting from a combination of high nuclear obliquity, high orbital eccentricity and an uneven distribution of active areas on the nucleus's surface¹³. Alternatively, Schwassmann–Wachmann 1 is at the right solar distance for its activity to be explained by amorphous–crystalline ice transitions, as suggested for comet Halley.

It is clear that there is still a great deal to be learned about cometary nuclei and the source or sources of their varied activity and behaviour. The glimpse afforded by the spacecraft flybys of Halley in 1986 was a great leap forward, but much more needs to be done. NASA is currently building a new spacecraft, CRAF (Comet Rendezvous Asteroid Flyby) for launch in 1997, to match orbits with short-period comet Kopff in late 2005¹⁴. CRAF will follow comet Kopff on its path around the Sun, observing the rise and fall of activity and gathering samples of coma gas and dust for onboard analysis. But budgetary pressures on NASA are high and the CRAF mission narrowly escaped cancellation by the US Congress in late September of this year. Given the interesting messages that Halley is still sending us from the outer Solar System, Congress has acted wisely by deciding in favour of this most fascinating scientific mission. □

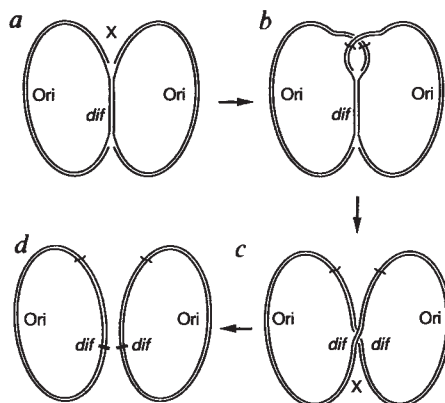
Paul Weissman is in the Earth and Space Sciences Division, Jet Propulsion Laboratory, Pasadena, California 91109, USA.

... and then there were two

Tania A. Baker

A HAPPY example of convergent evolution in research is to be found in three papers in *The New Biologist*^{1–3}. Together, the authors reveal the mechanism by which the circular chromosomes generally found in prokaryotes avoid becoming permanently joined to each other during genetic recombination (genetic exchange).

With circular chromosomes, the process of DNA replication is simplified as no special strategies are necessary to ensure synthesis of the chromosomal ends. But this elegant conformation comes at a price: the circular shape



The effect of homologous recombination between sister chromosomes during replication of a circular genome. *a*, An intermediate arising from bidirectional replication of a circular chromosome (Ori, the position of the replication origin; dif, a site within the replication terminus, located diametrically opposed to the origin). An odd number of genetic crossovers (shown by the X) between sister chromosomes in the replication intermediate results in the structure shown in *b* (~ is the site of the crossover). Completion of replication gives rise to a dimeric chromosome (*c*). Site-specific recombination by XerC at the dif sites can resolve this dimer into two monomeric chromosomes (*d*, // is the site of recombination between dif sites). If no homologous recombination or an even number of crossovers occur during replication, complete replication of the intermediate in *a* leads directly to the monomers shown in *d*, without the involvement of dif and XerC.

complicates the process of chromosome segregation. Circles frequently become physically joined during homologous recombination because an odd number of crossovers between two circles generates a dimer (see figure), and dimerization of sister chromosomes prevents normal chromosome partitioning into daughter cells during division. Clerget¹, Blakely *et al.*² and Kuempel *et al.*³ now provide convincing evidence that *Escherichia coli* encodes a site-specific recombinase which acts at a sequence within the

replication terminus. This recombination system apparently resolves into monomers the chromosome dimers that arise from homologous recombination.

Clues about the system operating in *E. coli* came from studies of naturally occurring plasmids (extrachromosomal genetic elements consisting of closed circles of double-stranded DNA). Low copy number plasmids (such as the F factor) encode, as part of their battery of inheritance functions, site-specific recombination systems that turn plasmid multimers into monomers. These recombinases are especially important for single-copy plasmids which require resolution of any dimers to ensure that each daughter cell acquires one plasmid. Higher copy number plasmids, such as ColE1, also carry information necessary to maintain them in a monomeric state⁴. In ColE1, a short DNA site called *cer* is the only plasmid sequence required for monomerization. The search for host genes required for *cer* function led to the discovery of the *xerC* gene in *E. coli*. The protein encoded by *xerC* is similar in sequence to the λ integrase family of site-specific recombinases and is responsible for monomerization of ColE1 (ref. 5).

Does XerC exist only to help plasmids survive? Evidently not. Clerget¹ found a chromosomal site of XerC action by studying another extrachromosomal element, the transposon Tn2350. This element is a component of the drug-resistance plasmid R1, but also occurs as a free circular form. A small sequence from Tn2350 is responsible for its unusual autonomy. This sequence also stabilizes replication-defective plasmids (that is, allows for the maintenance of the extrachromosomal element in dividing cells) and thus was originally thought to contain a replication origin. However, Clerget traced the stabilization activity of this fragment to its ability to recombine with the chromosome at high frequency. By integrating (and excising) very efficiently, DNA molecules lacking a functional origin are replicated along with the host cell chromosome. Clerget mapped the chromosomal site of Tn2350 integration to a unique locus within the *E. coli* replication terminus. Its sequence revealed a small region that is similar to both the Tn2350 stability sequence and to *cer*. Plasmid integration into this site required *xerC*, thus establishing it as a chromosomal target for the XerC recombinase.

Kuempel *et al.*³ independently identified the chromosomal locus of XerC action by a systematic deletion analysis

- Hainaut, O., Smette, A. & West, R. M. *IAU Circ. No.* 5189 (1991).
- Hughes, D. W. *Mon. Not. R. astr. Soc.* **251**, 26p–29p (1991).
- Gladman, B., Duncan, M. & Candy, J. *Celest. Mech.* (in the press).
- Duncan, M., Quinn, T. & Tremaine, S. *Astrophys. J.* **328**, L69–L73 (1988).
- Intrilligator, D. & Dryer, M. *Nature* **353**, 407–409 (1991).
- Krankowsky, D. in *Comets in the Post-Halley Era* (eds Newburn, R. L., Neugebauer, M. & Rahe, J.) 855–877 (Kluwer, Dordrecht, 1991).
- Prialnik, D. & Bar-Nun, A. *Astrophys. J.* **313**, 893–905 (1987).
- Herman, G. & Weissman, P. R. *Icarus* **69**, 314–328 (1987).
- Hartmann, W. K., Tholen, D. J., Meech, K. J. & Cruikshank, D. P. *Icarus* **83**, 1–15 (1989).
- Meech, K. J. & Belton, M. J. S. *Astr. J.* **100**, 1323–1338 (1990).
- Hahn, G. & Bailey, M. E. *Nature* **348**, 132–136 (1990).
- Jewitt, D. C. *Astrophys. J.* **351**, 277–286 (1991).
- Weissman, P. R. *Astr. Astrophys.* **187**, 873–878 (1987).
- Neugebauer, M. & Weissman, P. R. *EOS* **70**, 633 (1989).

of the replication terminus. The terminus region contains sequences that impede replication forks (*ter* sites) and genes involved in cell division. However, certain deletions that did not affect these known functions caused cells to accumulate abnormally long undivided forms. The locus responsible for this 'deletion-induced filamentation' (*dif*) was found to be the same as that used for integration of Tn2350.

Final confirmation that XerC promotes recombination at *dif* came from Blakely and colleagues², who inserted the sequence from *dif* that was most similar to the plasmid *cer*-like sites into a small test plasmid. The sequence promoted efficient *xerC*-dependent plasmid-plasmid recombination; and DNA-binding experiments using extracts from XerC-overproducing cells further indicated that the XerC protein specifically binds to both *dif* and *cer*.

Dramatic evidence for the importance of *xerC* and *dif* in *E. coli* cell division was revealed by microscopic observations of growing cell cultures. Strains deleted for either *xerC* (ref. 2) or *dif* (ref. 3) were viable but showed aberrant chromosomal partition and cell division. Cultures of either mutant contained filamentous cells and many cells apparently arrested just before cell division. Nucleoids sometimes appeared to contain unusually large amounts of DNA and were often incorrectly positioned within the cell. These morphological abnormalities were minimized in cells defective in homologous recombination, showing that recombination generates structures which must be resolved by XerC action at *dif* to allow normal cell division.

For recombination by XerC to be useful it must turn dimers into monomers without promoting the reverse reaction. How this one-way process is achieved at *dif* sites remains a mystery. Some other site-specific recombination systems, such as Tn3 resolvase, preferentially monomerize multimers because recombination sites on separate DNA molecules do not pair properly and thus fail to recombine⁶. Although a similar site-pairing mechanism for directional specificity may also operate at *dif*, this strategy may not be very efficient on DNA molecules as large as the *E. coli* chromosome. Indeed, even when *dif* sites are on plasmids, XerC promotes both plasmid multimerization and resolution. On the chromosome, where XerC presumably acts on the two *dif* sites at the termini of sister chromosomes, this lack of specificity would have potentially disastrous consequences. The signal essential for directional specificity is clearly absent from the *dif* plasmid substrates. What might serve as this signal in the chromosomal context?

Swift to avoid the swallow



EMPEROR penguins may be waddling jokes on land, but under water they can turn into regular rockets. This picture — a frame from the new IMAX film *Antarctica* now showing at several US museums — shows emperor penguins darting to and from a hole in the ice near their colony. Because they sense a predatory leopard seal nearby, the penguins are taking evasive action, accelerating from 0 to 7 m s⁻¹ in less than a second. By contrast, the seals that chase them rarely swim faster than 2 m s⁻¹, and so must rely on surprise to make a catch. **C.A.**

One possibility is that XerC might be able to take advantage of the partitioning of the chromosomes to 'tell' monomers from dimers. Given that *dif* is located within the replication terminus, XerC action might well be coupled to chromosome segregation. Perhaps the two *dif* sites pair and recombine whether they are on a dimeric or two monomeric chromosomes. Continual recombination would then repeatedly interconvert monomers and dimers. If the recombination machinery dissociates after each round, as the chromosomes are pulled apart the monomeric state may be 'trapped' when the two *dif* sites become spatially separated. This mechanism would often fail however, if the *dif* site is not the last point on the chromosome to be pulled apart. Alternatively, XerC recombination might hold the two *dif* sites together in a synaptic complex which is maintained while the rest of the DNA is untangled and separated. Resolution of the paired complex at *dif*, in the direc-

tion to give chromosome monomers, might then be triggered by the final segregational 'tug' on the DNA.

Little is known about the physical process of chromosome separation, so the validity of these models for XerC action cannot yet be tested. But the characteristics of the *xerC*-*dif* recombination system make it likely that in *E. coli*, as in eukaryotes, recombination is centrally involved in chromosome management during cell division. □

Tania A. Baker is at the National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA.

1. Clerget, M. *New Biol.* **3**, 780–788 (1991).
2. Blakely, G., Colloms, S., May, G., Burke, M. & Sherratt, D. *New Biol.* **3**, 789–798 (1991).
3. Kuempel, P. L., Henson, J. M., Dircks, L., Tecklenburg, M. & Lim, D. F. *New Biol.* **3**, 799–811 (1991).
4. Summers, D. K. & Sherratt, D. J. *Cell* **36**, 1097–1103 (1984).
5. Colloms, S. D., Sykora, P., Szatmari, G. & Sherratt, D. J. *J. Bact.* **172**, 6973–6980 (1990).
6. Bendnarz, A. L., Boocock, M. R. & Sherratt, D. J. *Genes Dev.* **4**, 2366–2375 (1990).

Greedy guts

REPPRESSED criminal inclinations are revealed in a study of the superficially innocent Anna's hummingbird, *Calypte anna* (L. Carpenter *et al. Evol. Ecol.* **5**, 405–414; 1991). At least as indicated by the size of the crops in which they store foraged nectar, these birds have a capacity for larger meals than required for energetically efficient feeding. The new experiments, using artificial feeders, confirm that territorial males take frequent, small meals as expected. But when poaching from another's territory, they sacrifice airworthiness for feeding binges that are snatched while the owners are not looking. The enlarged crops apparently are designed for this antisocial behaviour.

Kerb crawling

THE Hayward fault, across the bay from the San Andreas fault, could be due for a powerful earthquake, according to J. J. Lienkaemper *et al. (J. geophys. Res.* **96**, 18,261–18,283; 1991). The site of two earthquakes in 1836 and 1868, the 82-km fault is now considered the most likely site for the next severe Californian earthquake following that at Loma Prieta in 1989. Lienkaemper *et al.* enlisted the help of local city departments to assess the scale of threat posed. They selected long, initially straight constructions perpendicular to the fault (road kerbs, fences, a water tunnel) to look for any misalignment caused by creep. From the ages supplied by the cities they derived the local creep rate. The fastest creep, about 9 mm a year at the southern end, the authors infer to be the underlying fault movement. From the deficit elsewhere, they suggest that, should the whole fault fail in a single go, an earthquake not much less powerful than 1989's could be possible.

Small wonder

INTEGRATION in electronic circuits nowadays extends only as far as the power source. Then (unless they use Sun-driven solar cells) electronic engineers rely on external sources for power, such as button-sized batteries. But F. K. Shokooki *et al. (Appl. Phys. Lett.* **59**, 1260–1262; 1991) have brought the prospect of true integration closer with the development of electrodes to be used in micro-batteries. Made of manganese oxide, the cathodes operate by absorbing lithium ions into their porous interiors from a lithium-chlorate-based electrolyte. Cells made with micrometre-thick cathodes and thin metallic anodes generate a voltage of 4 volts, currents of a few tens of microamps per square centimetre, and (most importantly) can be discharged and recharged up to 70 times before deterioration shows. The next step is to incorporate the batteries into a chip.

Chaos in ocean heat transport

Curt Covey

THE search in the Earth's temperature record for proof of human amplification of greenhouse warming has so far been inconclusive, perhaps because of obfuscation of the signal by climatic 'noise' over timescales of one or more decades. On page 836 of this issue¹, Weaver *et al.* add to an impressive body of evidence that much of this noise comes from the oceans. They find that the thermohaline circulation, in which warm water flows polewards at the surface while colder water returns to lower latitudes at depth, can sustain oscillations between very different states. Resulting variations in the oceans' transport of heat from the warm tropics to the cooler poles are substantial and often chaotic, according to the authors' calculations.

The character of thermohaline circulation depends on where surface water sinks to form bottom water, a process sensitive to water's density as determined by temperature and salt concentration (hence the name). The latter effect dominates in the sinking of polar water and, unlike temperature, is not subject to local feedback; the salinity of surface water is established by the difference between evaporation and precipitation ($P-E$), neither of which are affected directly by surface salinity. So it is possible, as suggested a decade ago by Rooth², that there are different quasi-stable states of circulation and transitions among them. Subsequent calculations on one-dimensional³ and two-dimensional⁴ ocean-circulation models, in addition to three-dimensional modelling acknowledged by Weaver *et al.*, have verified the plausibility of Rooth's conjecture. These studies have also focused on the North Atlantic, where much of the oceans' formation of bottom water takes place — in the present state of the climate system, at least.

The new contribution by Weaver *et al.* is to explore in unprecedented detail the history of transitions between the different states of circulation, and the way these transitions depend on the spatial distribution of $P-E$. Doing so requires a bold modelling tactic. To speed up their three-dimensional numerical simulation of the oceans, the authors not only restrict its domain to the North Atlantic and simplify its geometry (both standard techniques), but also linearize the fluid-dynamical equations that enforce conservation of horizontal momentum. Thus the most numerically irritating non-linearity, in which fluid velocity transports fluid velocity itself, is apparently eliminated, one hopes without com-

promising the transport of heat.

Weaver *et al.* promise that details of their method will be published soon; as best I can guess, it resembles the 'large-scale geostrophic' assumption in Maier-Reimer's global circulation model⁵. As the name suggests, Maier-Reimer's technique is based on the reasonable assumption that the oceans are close to geostrophic balance — that is, pressure gradients are nearly balanced by Coriolis forces. Because Coriolis forces depend linearly on velocity, the fluid velocity is determined without reference to non-linear transport. Consequently, the model of Weaver *et al.* is capable of stunningly long integrations, spanning more than 10,000 years of simulated time.

Thus equipped, Weaver *et al.* begin their simulations with traditional (albeit unphysical) boundary conditions restoring surface salinity to today's observed values. Under these conditions the model reaches a steady state in which bottom water forms at latitudes closest to the North Pole. The $P-E$ fluxes that would be required to maintain this state are then imposed as a new condition. As in many previous studies, the circulation established with the old boundary condition proves to be unstable with the $P-E$ flux boundary condition.

What is new is chaotic oscillation between the old steady state and a new circulation pattern in which bottom water formation occurs at 54° N instead of at the Pole. A local maximum in $P-E$ at this latitude makes surface water more saline, hence more dense, and sinking occurs. Weaver *et al.* find this change is itself unstable: the entire thermohaline circulation speeds up, surface water "passes rapidly through the evaporative region and hence the surface waters do not become as saline". The circulation returns to sinking at the Pole, and the process repeats perhaps a thousand times over nearly 8,000 years.

Weaver *et al.* show that their chaotic oscillations depend on the local $P-E$ maximum by repeating their experiment with artificially altered $P-E$ fluxes in which this maximum is removed. The oscillations disappear, but whether there are one or two stable states of circulation depends on the details of the $P-E$ distribution. Perhaps the authors have

1. Weaver, A. J., Sarachick, E. S. & Marotzke, J. *Nature* **353**, 836–838 (1991).
2. Rooth, C. *Prog. Oceanog.* **11**, 131–149 (1982).
3. Gaffin, S. R., Hoffert, M. I. & Volk, T. J. *geophys. Res.* **91**, 3944–3950 (1986).
4. Stocker, T. F. & Wright, D. G. *Nature* **351**, 729 (1991).
5. Bacastow, R. & Maier-Reimer, E. *Climate Dyn.* **4**, 95–125 (1990).
6. Watts, R. G. & Morantine, M. C. *Climatic Change* **18**, iii–vi (1991).

sampled the typical transition of non-linear systems, in which one stable state bifurcates to two quasistable states, then four and so on until 'order' becomes 'chaos'. If so, this work has uncovered more than a mathematical nicety. Watts and Morant⁶ directly show that variations of thermohaline circulation intensity of a few per cent could account for the most vexing part of this century's climate noise. Their paper is appropriately entitled, "Is the Greenhouse Gas—CRYSTALS

Climate Signal Hiding in the Deep Ocean?". In the light of Weaver *et al.*'s findings, one might add the question: does the chaotic nature of oceanic circulation limit the predictability of climate, just as the chaotic nature of atmospheric circulation is known to limit the predictability of weather? □

Curt Covey is at the Lawrence-Livermore National Laboratory, Mail Stop L-264, Livermore, California 94551, USA.

Surfaces make a difference

Richard J. Reeder

CRYSTALS acquire isotopic heterogeneities during growth as a result of differences in surface structure, Dickson reports on page 842 of this issue¹. Dickson shows that differential fractionation of carbon and oxygen isotopes occurs at structurally nonequivalent faces of natural calcite (CaCO_3) crystals, introducing inhomogeneities into the bulk crystals. This fractionation difference reflects disequilibrium, and will therefore be noted with interest by geochemists, who often rely on isotope data but assume that fractionation is an equilibrium process.

Dickson's micro-sampling and mass-spectrometric study focuses on inorganic calcite crystals that exhibit several nonequivalent faces. Within individual crystals, systematic differences of up to 2‰ for the isotope ratio $^{13}\text{C}/^{12}\text{C}$ and nearly 1‰ for $^{18}\text{O}/^{16}\text{O}$ are clearly revealed between regions ('growth sectors') in the crystal that grew simultaneously, but from nonequivalent crystal faces. It is essential that any isotope differences are shown to occur between time-equivalent growth sectors (or time-equivalent portions of them), because the isotope composition of the growth solution may change with time.

This spatial distribution of stable isotopes is similar in every respect to the sector zoning, involving differences of elemental abundance, that is commonly seen in mineral crystals (also known as anomalous or anisotropic segregation in industrial crystallization). Indeed, one of Dickson's calcite crystals displays such sector zoning. It is now evident that this face-dependent stable-isotope fractionation is but another member of a much broader class of surface-mediated growth phenomena capable of introducing heterogeneities into the crystal. Not only are variations of isotope and elemental chemistry possible, but also ordering and local reduction of symmetry can occur. And these seemingly anomalous growth phenomena occur over a broad spectrum of crystalline materials representing a

considerable range of natural and laboratory growth environments.

Sector zoning found in mineral crystals is perhaps the best studied case. The chemical differences between sectors commonly involve minor or trace constituents, such as Mn^{2+} in calcite or Fe^{3+} in quartz, but may also involve major elements, Si and Al in the silicate mineral augite², for example. Face-dependent partitioning differences of trace elements are also well known in semiconductor growth. Most interpretations have attributed the differential incorporation to the differences in surface structure of nonequivalent faces.

Recent work on synthetic diamond and on calcite reveals even more highly organized surface controls on trace-element incorporation resulting from spiral growth. Crystals that grow by layer-spreading mechanisms commonly rely on growth steps that originate at dislocations emerging from the crystal. These steps may form either a spiral pattern or, where step segments are mostly straight, a polygonal pattern (Fig. 1). For the latter case, each growth spiral hillock is composed of slightly inclined flanks, or vicinal faces. Frank *et al.*³ demonstrated last year that nitrogen is incorporated onto nonequivalent vicinal faces of diamond with systematic differences. The same has been found for Mg, Sr, Mn^{2+} and Fe^{2+} incorporation on individual calcite faces⁴. It seems, then, that the detailed structure of the growth steps, which varies with orientation on a face, is responsible for differential incorporation.

Others have shown that growth steps, which virtually always have less sym-

metry than the bulk crystal, may induce ordering of atoms as the steps spread laterally during growth. The result can be a lowering of symmetry in different regions of the crystal, again, typically associated with vicinal faces of growth hillocks. Such lowering of symmetry may give rise to anomalous optical properties and is commonly attended by twinning.

The face-dependent, stable-isotope fractionation observed by Dickson for calcite is by all appearances similar, and not entirely unexpected. In fact, an isotope-fractionation difference of greater than 40‰ between cubic and octahedral growth sectors has been seen with nitrogen impurities in synthetic diamond⁵. Perhaps the most significant contribution of the isotope work will be new insight into the mechanisms by which the surface selectivity operates. For all of these cases, the chemical, isotope and structural inhomogeneities are related to surface structural differences, either between different faces or within a single face. As the surfaces translate forward during growth, surface-controlled inhomogeneities are incorporated into the crystal, after which they may be difficult or impossible to modify (Fig. 2). These observations indicate a kinetic or mechanistic control on



FIG. 1 A large growth hillock seen on the face of a calcite crystal is composed of four shallow vicinal faces, each coloured differently in this differential-interference contrast image owing to very slightly different inclinations. Within each vicinal face, growth steps have a unique orientation, as may be seen in the two smaller vicinals at upper right. Trace elements are partitioned differently between the vicinal faces. In other types of crystals, differences of ordering occur because of different surface structures. Are stable isotopes similarly affected?

incorporation, rather than an equilibrium thermodynamic one.

The primary factors determining specific selectivity for chemical and structural anomalies presumably include ion (or molecule) size, charge and shape. For isotope fractionation differences, however, mass effects associated with differences in vibrational frequencies must operate. This implies that surfaces, or

steps, provide sites for incorporation that differ somehow in bonding characteristics. Or perhaps some coupled effect may be involved. Stable C and O isotope fractionation is known to vary with chemical composition in some carbonate minerals. One wonders if the sector-related isotope differences might be following chemical differences.

The observation of disequilibrium C and O isotope fractionation in calcite may have important implications for car-

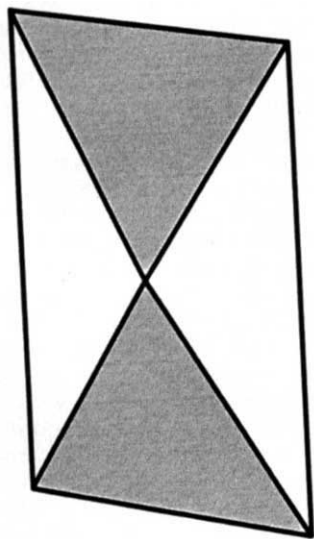


FIG. 2 A slice through an idealized crystal bounded by two nonequivalent forms with their associated growth sectors (shaded differently). Isotope, chemical and structural differences between sectors result from dissimilar surface structure on nonequivalent faces during growth.

bonate geochemists, who have relied on differences and trends in isotope data from limestones and other calcite-bearing rocks, assuming equilibrium fractionation has occurred. Although Dickson's observed fractionation differences are not extremely large, his findings should cause geochemists to question this basic assumption. Additionally, it is evident that many current studies involving carbonate minerals place greater emphasis on spatially and temporally constrained microsampling techniques. Small differences in isotope composition that are sometimes interpreted as variations of temperature, or of fluid isotope composition, may need to be more carefully evaluated. □

Richard J. Reeder is in the Department of Earth and Space Sciences, State University of New York, Stony Brook, New York 11794, USA.

Mapping the way forward

Kevin Davies

HUMAN genetics has come a long way in the thirty-five years since the first International Congress of Human Genetics in Copenhagen in 1956, when Tijo and Levan announced that there are 46 chromosomes in a cell, not 48 as some had believed. More recently, the advent of genetic and physical mapping approaches has brought the isolation of various human disease genes within reach, in many instances providing the only realistic route towards defining the primary biochemical defect. To judge from the eighth quinquennial conference*, the pace at which the genetic basis of simple and complex diseases is determined continues to quicken.

Many important disease genes have been identified even in the past few months. Some have been isolated without prior knowledge of their structure, by constructing detailed physical maps around closely linked markers, so-called positional cloning (Kallmann's^{1,2} and fragile X syndromes³, colorectal cancer⁴). Others have been characterized by using a candidate gene approach (early-onset familial Alzheimer's disease⁵⁻⁷ and Marfan syndrome⁸).

Last year, the latter strategy led to the discovery of mutations in the rhodopsin gene⁹ as the cause of an estimated one quarter of cases of autosomal dominant retinitis pigmentosa (ADRP) — progressive blindness which affects about 1 in 4,000 people and is known to have several genetic causes. P. Humphries (Trinity College, Dublin) and T. Dryja (Harvard Medical School, Boston) have identified mutations in a second gene for ADRP; this gene, located on chromosome 6, codes for a protein called peripherin, and is the human homologue of the *rdx* (retinal degeneration slow) gene which is defective in a mouse model of blindness^{10,11}. Furthermore, a putative mutation has been found in a third gene, *ROM1* (R. McInnes, Hospital for Sick Children, Toronto), which maps to chromosome 11q13 and encodes a protein that is 35 per cent identical to peripherin and probably plays an important developmental and structural role in the rod photoreceptor. A fourth ADRP locus has been mapped by linkage in a large Kentucky family to chromosome 8 (S. Daiger, University of Texas Medical School, Houston; ref. 12).

A recent success of positional cloning was the isolation of the gene for choroideremia¹³, a common sex-linked form of hereditary retinal degeneration. F. Cremers (University of Nijmegen,

Netherlands) described the isolation of a highly related gene which, on the basis of its localization to the long arm of chromosome 1, is now an attractive candidate gene for one form of Usher syndrome (type II), in which relatively late-onset retinitis pigmentosa is accompanied by hearing loss. An adult-onset deafness gene was mapped to chromosome 5q by linkage in a large Costa Rican pedigree, descendants of a man who lived in the eighteenth century (M.-C. King, Stanford University).

With one notable exception, that of cystic fibrosis (CF), the positional cloning of disease genes has been expedited by mapping chromosomal rearrangements or large-scale deletions in affected patients, as seen most recently for the neuronal migration disorder, Kallmann's syndrome^{1,2}. Indeed, the laborious progress in isolating the defective genes for Huntington's disease, myotonic dystrophy and polycystic kidney disease, for example, which were mapped several years ago, can be attributed to the failure to observe such rearrangements. Analysis of recombination positions, haplotypes and linkage disequilibrium are helping to compress the candidate regions, with most success for myotonic dystrophy, which appears to lie in a 250-kilobase interval on chromosome 19 (R. Korneluk, University of Ottawa, Ontario). By contrast, the gene for Huntington's disease still cannot be positioned with confidence to much less than 2 megabases near the tip of chromosome 4 (G. Bates, Imperial Cancer Research Fund, London).

As if to underline the importance of chromosomal rearrangements in defining disease genes, preliminary findings indicate the Lowe oculocerebrorenal syndrome gene has been identified and cloned (I. Okabe, University of Pennsylvania). Lowe syndrome is an X-linked disorder resulting in congenital cataracts, mental retardation and renal-tubule abnormalities. The candidate gene spans an Xq25 translocation breakpoint in one patient with Lowe syndrome, and the RNA transcript is either absent or of abnormal size in several others. Prader-Willi (PWS) and Angelman (AS) syndromes differ considerably in their symptoms yet both result from cytogenetic deletions of similar regions of the paternal and maternal chromosome 15 long arm, respectively, and thus provide some of the best evidence to date for genomic imprinting in humans. But it is not clear how many, or even if the same, genes contribute to both disorders¹⁴. M. Lalande (Harvard Medical School,

1. Dickson, J. A. D. *Nature* **353**, 842–844 (1991).
2. Hollister, L. S. & Gancarz, A. J. *Am. Mineral.* **56**, 959–979 (1971).
3. Frank, F. C. et al. *J. Crystal Growth* **100**, 354–376 (1990).
4. Paquette, J. & Reeder, R. J. *Geology* **18**, 1244–1247 (1990).
5. Boyd, S. R., Pillinger, C. T., Milledge, H. J., Mendelsohn, M. J. & Seal, M. *Nature* **331**, 604–607 (1988).

*8th International Congress of Human Genetics, Washington DC, 7–11 October, 1991.

Boston) described a three-generation Japanese family segregating for a sub-microscopic 15q deletion which shows no associated phenotype from grandfather to daughter, yet all three grandchildren have AS. The lack of PWS in the mother provides some of the best evidence so far that different genes are involved in the two syndromes.

For most of the major human disease genes already isolated, the correlation between genotype and phenotype is a priority; thus the findings described by J. Amos (Boston University School of Medicine) surprised many. Congenital bilateral absence of the vas deferens (CBAVD) had been regarded as a distinct mendelian disorder, but upon screening alleles of the cystic fibrosis transmembrane regulator (CFTR) gene from 25 CBAVD patients for five common CF mutations, 60 per cent of patients were found to have at least one mutant CF allele. Of greater significance, novel mutations of the second allele were seen in three of the patients, who otherwise had no apparent signs of cystic fibrosis. Thus, compound heterozygosity of CFTR mutations seems capable of producing a very mild form of the disease, with important implications for diagnosis and family counselling.

The phenotypic spectrum of dystrophin mutations may also be wider than previously thought. Duchenne and Becker muscular dystrophies are well-known to result from mutations or deletions in the dystrophin gene, but J. Towbin (Baylor College of Medicine) showed that X-linked dilated cardiomyopathy appears to be due to a specific abnormality in the cardiac, as opposed to skeletal, muscle expression of dystrophin.

A few predictions, some specific and others more general, are likely to be met by the time of the ninth Congress in 1996. For example, a couple of rapidly evolving techniques may have a profound impact on prenatal diagnosis and DNA fingerprinting. D. Bianchi (Children's Hospital, Boston) and S. Wachtel (University of Tennessee) described the

first successful screening of chromosomal trisomies in fetal cells separated from maternal blood during early pregnancy. The task ahead is to improve the sensitivity and reliability of the technique. And A. Jeffreys (University of Leicester) introduced a new PCR-based analysis of a hypervariable repeat sequence that can be translated into a simple yet extraordinarily informative digital code, which will have substantial implications for the study of human populations and, of course, for forensic science¹⁵.

In more general terms, positional cloning, aided by improved techniques for identifying gene sequences in large

stretches of DNA, could yield over 100 new disease genes in the next five years or so (F. Collins, University of Michigan). And new linkage approaches in rodent models for complex polygenic traits will pave the way for identifying genes for diabetes¹⁶, hypertension^{17,18} and epilepsy¹⁹, to name but a few (E. Lander, Whitehead Institute, Cambridge). As if to meet demand, at least two promising new human genetics journals will be launched next year, while others are destined for a facelift. The extra reading should be worthwhile. □

Kevin Davies is an assistant editor of *Nature*.

Most of the senses begin to make some sense

Doron Lancet

SENSORY cells convert different forms of energy to the brain's 'currency' — transmembrane potentials. But they do more. They have mechanisms for self-modulation, and their built-in diversity generates the complex patterns that underlie perceptual images. At a meeting* held last month, a unified view emerged of how this is accomplished in photoreception, mechanoreception and chemoreception.

Colour vision is based on multiplexing three cone detectors of different spectral sensitivities¹. Contrary to predictions, neutral, not charged residues mainly determine at which wavelength each visual pigment absorbs best (J. Nathans, Johns Hopkins University). The absorption of a single photon results in a considerable energy change relative to thermal noise (S. Block, Rowland Institute). Specific glutamate and histidine residues may internally orchestrate (perhaps by a serine protease-like mechanism) the use of the right energy fraction to generate a conformation capable of activating the GTP-binding protein (J. Nathans).

That GTP-binding proteins (G-proteins) serve as cellular photo-amplifiers² inspired a search for a similar pathway in olfaction^{3,4}, which culminated in the cloning of candidate seven-transmembrane-domain receptors for odorants⁵. New evidence for the receptors' functional identity is provided

by a specific antibody that strongly stains the ciliary extensions of the sensory neurons (R. Reed, Johns Hopkins University). How the diverse olfactory receptors, which are at least 100 times more numerous in kind than colour vision pigments, contrive to produce an odour image, is still an open question. The observed activity patterns in olfactory bulb glomeruli might arise if axon targeting to a glomerulus was correlated with the type of olfactory receptor ex-

1. Franco, B. *et al.* *Nature* **353**, 529–536 (1991).
2. Legouis, R. *et al.* *Cell* **67**, 423–436 (1991).
3. Kremer, E. J. *et al.* *Science* **252**, 1711–1714 (1991).
4. Kinzler, K. W. *et al.* *Science* **253**, 661–665 (1991).
5. Goate, A. *et al.* *Nature* **349**, 704–706 (1991).
6. Murrell, J., Farlow, M., Ghatti, B. & Benson, M. D. *Science* **254**, 97–99 (1991).
7. Chartier-Harlin, M.-C. *et al.* *Nature* **353**, 844–846 (1991).
8. Dietz, H. C. *et al.* *Nature* **352**, 337–339 (1991).
9. Dryja, T. P. *et al.* *Nature* **343**, 364–366 (1990).
10. Farrar, G. J. *et al.* *Nature* (in the press).
11. Kajiwara, K. *et al.* *Nature* (in the press).
12. Blanton, S. H. *et al.* *Genomics* (in the press).
13. Cremers, F. P. *et al.* *Nature* **347**, 674–677 (1990).
14. Wagstaff, J. *et al.* *Am. Soc. hum. Genet.* **49**, 330–337 (1991).
15. Jeffreys, A. J. *et al.* *Nature* (in the press).
16. Todd, J. A. *et al.* *Nature* **351**, 542–546 (1991).
17. Hilbert, P. *et al.* *Nature* **353**, 521–529 (1991).
18. Jacob, H. J. *et al.* *Cell* **67**, 213–224 (1991).
19. Rise, M. L., Frankel, W. N., Coffin, J. M. & Seyfried, T. N. *Science* **253**, 669–673 (1991).

*Sensory Transduction, Marine Biological Laboratory, Woods Hole, Massachusetts, 5–8 September 1991.

	Vision	Olfaction	Taste	Audition
Detection	←	Protein receptors	→	Organelle movement
Signal modulation	Filters, lenses	← Enzymes, carriers →		Outer hair-cell feedback
Quality coding	Absorption spectra	← Ligand specificity →		Vibrational resonance
Sensitivity (amplification)	←	G-proteins, second messengers	→	Direct link to channels
Cellular response	← Ion-channel modulation →			
Calcium-related adaptation	Guanylate cyclase activation	Blocking of cAMP channels	?	Sliding motor adjustment

There are similarities and differences at each level from sensory detection to coding and perception. Signal modulation occurs before interaction with the detector, and in olfaction it may involve biotransformation enzymes and odorant-binding proteins⁸. The prevailing principle in quality coding is the multiplexing of many detectors with different but often broad selectivities^{1,4,6,7}. There are adaptation and modulation mechanisms other than those shown^{2,4,7,8}, and in the case of taste they await full description.

pressed by the primary sensory neuron⁴ (G. Shepherd, Yale University). The fascinating story of organization and differential expression in the 'olfactory sub-genome' should soon unfold: there may be a thousand intronless receptor-coding regions, possibly clustered, and perhaps occupying over ten megabases (L. Buck, Columbia University; R. Reed).

With fewer ligands (and possibly re-

ceptors) to cope with, taste curiously uses many more transduction devices⁶, but none had been identified by cloning until very recently. The characterization of a taste-bud-specific G-protein (G_{gust}) of the G_i/G_o subfamily (R. Margolskee, Roche Institute) breaks this stalemate, and we can now expect an avalanche of molecular characterizations. S. Kinnamon (Colorado State University) produced the first direct electrophysiological support for the idea that sweet responses involve cyclic AMP, hence a G_s -like protein. So G_{gust} is more likely to be related to bitter taste (the sour and salty tastes involve direct action on cation channels). These results, together with the rather well-defined sweet and bitter ligand 'pharmacology' and the potential use of a congenic bitter non-taster mouse (A. Spielman, New York University) mean that gustation may well soon disclose its own seven-transmembrane-domain receptors.

Audition's counterpart of diverse protein detectors is the organ of Corti with its spatially segregated acoustic frequency responses⁷. Instead of using G-proteins and second messengers to achieve high sensitivity, the cochlear hair cells employ an ion-channel gating 'lid' directly hooked to a pivoting stereocilium. The connector is a 'tip-link' about 100 nm long, whose dissociation at low levels of calcium disrupts the auditory mechanism (D. Corey, Harvard University). A displacement smaller than 0.1 nm (near to thermal noise) then suffices to effect channel opening. The channels (about one for each stereocilium) reside near the stereociliary tips, and channel opening astonishingly yields a discernible dip in the stiffness-displacement curve for these organelles⁷, further evidence for a direct link (J. Hudspeth, University of Texas).

Sensory signals are modulated before interacting with the receptive apparatus (see the table on the previous page). A prime example is the cochlear outer hair cell, which serves as a positive feedback preamplifier for incoming acoustic waves. It has a voltage-driven cytoskeletal motor embedded in the membrane (P. Dallos, Northwestern University; J. Ashmore, School of Medical Sciences, Bristol). In video recordings, outer hair cells 'dance' visibly to an electric potential tune, whose *in vivo* counterpart is played by efferent neurons.

An emerging similarity in adaptation mechanisms (see table) is that many of them involve calcium ions. Calcium depletion in vertebrate rods, resulting from closure of cGMP-gated channels, activates a guanylate cyclase, thus returning cGMP to a higher level². In invertebrates, the story is still unfolding: calcium, via the *trp* protein, may help maintain excitation (in flies, B. Minke,

Hebrew University, Jerusalem), and cGMP channel-gating could underlie transduction after all (in *Limulus*, J. Lisman, Brandeis University). There is a similar cation channel in olfactory cilia, curiously bearing a sequence hallmark of a cGMP-binding protein (B. Kaupp, Institute for Biological Information, Jülich); yet functionally it responds to odorant-elevated cAMP. The olfactory channels are blocked by an increase in intracellular calcium, and this is a possible mechanism for adaptation of the odorant response (S. Firestein, Yale University) in addition to adaptation by cAMP phosphodiesterase and possibly cAMP-dependent kinase (H. Breer, Stuttgart University).

Auditory cells have unusual adaptation mechanisms and, again, they are affected by changes in intracellular calcium (R. Fettiplace, University of Wisconsin). Bacterial chemotaxis is an example of near-perfect adaptation, in that the response is purely to a change (positive or negative) of chemoattractant concentration over a wide range of initial conditions (H. Berg, Harvard University). But how is this achieved when a positive signal means stretching the tip-link between two stereocilia, and a negative one means loosening it? A rewarding answer is that a sliding filament motor could allow one attachment point to slide back and forth on a long track, and still keep a standing tension. In a *tour de force* of experiments and mathematical analysis, D. Corey provided evidence for this hypothesis. An actomyosin-like mechanism is the prime candidate because of similarities in the biophysical characteristics and because stereocilia are chock full of actin filaments. Homology-based cloning strategies directed at actin and myosin could help in the formidable task of identifying the auditory molecules, perhaps present at one copy for each organelle.

Our anthropocentric bent means that responses to light, sound and flavour are being tackled first. Thus, for example, the magnetosome, a magnetoreception organ in salmon and bees (but less obviously in humans) is still being sought (J. Kirschvink, Caltech). My hope is that this and other such enigmatic modalities will soon also take centre-stage in sensory research. □

Doron Lancet is in the Department of Membrane Research and Biophysics, The Weizmann Institute of Science, Rehovot, 76100 Israel.

1. Nathans, J. *A. Rev. Neurosci.* **10**, 163–194 (1987).
2. Stryer, L. *J. biol. Chem.* **266**, 10711–10714 (1991).
3. Reed, R. R. *Cell* **60**, 1–2 (1990).
4. Lancet, D. *Nature* **351**, 275–276 (1991).
5. Buck, L. & Axel, R. *Cell* **65**, 175–187 (1991).
6. Kinnamon, S. C. *Trends Neurosci.* **11**, 491–496 (1988).
7. Hudspeth, A. J. *Nature* **341**, 397–404 (1989).
8. Burchell, B. *Nature* **350**, 16–17 (1991).

Personal vibes

IDENTIFYING yourself to strangers or machines is a problem these days. Signatures can be forged, keys and machine-readable cards can be stolen. What is needed, says Daedalus, is some way of being machine-readable yourself.

He reckons that each of us has a unique physical style or 'body language'. Paradoxically, it is revealed most clearly by the act of doing nothing. If you try to hold your hand quite still, for example, it will actually execute a tiny, continuous, complex stochastic three-dimensional tremor. This may seem like random positional noise, but its statistical constants — its velocity power-spectra in each dimension and the phase-relationships between its main components — form a 'signature' unique to your own neuromuscular system. No mimic could copy it, even if he knew what to copy, for he could not eliminate his own characteristic tremor. But modern, fast-response capacitive probes, sensitive to movement in microns and frequencies in kilohertz, could record the signature in great detail. Daedalus's new ID system features a cavity lined with such sensors and coupled to a computer.

To identify yourself to this system, you merely need to put your hand (or maybe just a finger) in the sensor cavity and hold it as still as possible for a few seconds. The system rapidly analyses the microscopic quivering motion and derives your neuromuscular-tremor signature, or 'Tremsign'. If this matches your stored Tremsign, the system then opens the door, hands over the money, or whatever. It can even write some coded transform of the captured signature on a document (a traveller's cheque, say). Only you will be able to reproduce that signature on another terminal.

At first Daedalus worried that your Tremsign would vary greatly with fatigue, or recent manual labour, or even an attack of shivering or 'flu. But he now argues that you even shiver in your own body language, so to speak. A good Tremsign algorithm would still identify you, and might also identify the cause of the perturbed signature. So Tremsigns could turn out to be useful for medical or psychiatric diagnosis as well as for identification. Thus a Tremsign-protected ballistic-missile installation, or even a car, might refuse to be activated if it recognized its authorized operator, but reckoned he was drunk or crazy.

Tremsigns will wonderfully reduce our dependence on ID cards and other insecure identifiers. And the growing army of illiterates, barred from full effective citizenship because they cannot sign their names, will have one less barrier to overcome.

DAVID JONES

Binary precursor for planet?

SIR — Bailes *et al.*¹ have recently discovered a pulsar, PSR 1829–10, which shows time-delay variations, indicating that it is orbited by an object of planet mass. It is surprising both that a planet has apparently survived the supernova explosion that formed the neutron star and that the planet is now in a circular six-month orbit. We have devised a plausible explanation, in which the neutron star is relatively old and was in a wide orbit around a massive star. Be X-ray binaries such as this are relatively common². When the companion star swelled up to become a giant, it transferred mass to the neutron star which thereby acquired a massive disk. We assume that the disk both spun up the neutron star and, after the supernova explosion of the companion, that the outer parts formed planets, at least one of which is massive enough to cause the observed time-delay variations.

To present our argument in more detail: we require that the progenitor system was an X-ray binary, for example a Be X-ray binary, with a relatively long orbital period of at least 5 yr. This means that the neutron star is not captured by frictional drag in the envelope of its companion when it becomes a red giant (it escapes 'spiral-in'). The neutron star may have been spun down to a spin period of minutes due to the propeller effect³ in the matter expelled by the companion in its wind and equatorial disk. When the massive companion does swell to become a giant, the wind increases in strength with $\dot{M}_{\text{wind}} \sim 10^{-5} M_{\odot} \text{ yr}^{-1}$ and also slows down. If the wind velocity is less than the orbital velocity of the neutron star, it will accrete a fraction of $\dot{M}_{\text{wind}}$ equal to the square of the mass ratio of the stars, which is about 1 per cent in the case considered (where the mass of the giant is about 10 times that of the neutron star). This implied accretion rate is super-eddington. The material cannot all be accepted by the neutron star, so a large disk of mass $> 10^{-3} M_{\odot}$ accumulates around it provided that this phase lasts longer than 10^4 yr. The disk extends out to the roche lobe, which means that its outer radius is several AU.

When the companion explodes in a supernova, the binary system becomes unbound. Most of the disk will survive if it subtends a small angle at the companion or is shielded from objecta in the orbital plane. It then resembles the inner parts of the early proto-solar system. The central part will accrete onto the neutron star, thereby spinning it up until it reaches an equilibrium with the keplerian rotation of the disk at the magnetospheric radius. Our chief assumption

now is that the outer parts of the disk cool and form planets. This should occur after the luminous companion has exploded if the inner parts shield the outer parts from the X-rays. By analogy with our own Solar System, we suppose that it is reasonable for the planet masses to be similar to or greater than that of the Earth. The planets will, of course, have orbits of low eccentricity.

The observed large spin period derivative of the pulsar indicates that it lies above the classical spin-up line for neutron stars⁴, which could argue against our model. We note that this line, which is theoretically derived from (uncertain) models of the disk-magnetosphere interaction at the eddington limit, does not necessarily apply to a high magnetic field neutron star such as is considered here. (For instance, the neutron star may accrete at a super-eddington rate like the $10^{39} \text{ erg s}^{-1}$ X-ray pulsar, AO538–66, in the Large Magellanic Cloud⁵ since it has a strong magnetic field which can channel the accretion flow and beam the outgoing radiation.) We also note that a similar scheme to ours may be relevant to the X-ray pulsar 1E2259+586 (ref. 6 and references therein) which is in a supernova remnant and appears to be powered by accretion yet has no measured binary motion. It has a period of nearly 7 s, which is much larger than the 0.3-s period of PSR 1829–10. In our scheme it may just have had a much shorter spin-up phase than PSR 1829–10 and is now a single object surrounded by, and accreting from, a massive disk.

The advantage of our scheme is that it involves only a simple extrapolation from known systems and it does not require a sequence of improbable events. Planetary systems around stars do not have to be common. We predict that there should be another neutron star (from the ex-Be companion) near PSR 1829–10 and possibly an old supernova remnant (this depends on when the companion exploded). We also expect that PSR 1829–10 has further planets also in circular orbits.

A. C. FABIAN
Ph. PODSIADLOWSKI

*Institute of Astronomy,
Madingley Road,
Cambridge CB3 0HA, UK*

1. Bailes, M., Lyne, A. G. & Shemar, S. L. *Nature* **352**, 311–313 (1991).
2. van den Heuvel, E. P. J. & Rappaport, S. A. in *Physics of Be Stars* (eds A. Slettebak & T. P. Snow) 291–307 (Cambridge Univ. Press, 1987).
3. Illarionov, A. F. & Sunyaev, R. A. *Astr. Astrophys.* **39**, 185–195 (1975).
4. Battacharya, D. & van den Heuvel, E. P. J. *Phys. Rep.* **203**, 1–124 (1991).
5. Skinner, G. K. *et al.* *Nature* **297**, 568–570 (1982).
6. Davies, S. R., Wood, K. S. & Coe, M. J. *Mon. Not. R. Astr. Soc.* **245**, 268–274 (1990).

■ See also Letters on pages 827 and 829.

CJD discrepancy

SIR — Palmer *et al.*¹ report that homozygosity, Met/Met or Val/Val, at codon 129 of the prion protein (PrP) confers genetic predisposition to sporadic Creutzfeldt–Jakob disease (CJD) in the United Kingdom. But this is not true in Japanese patients with sporadic CJD, where either homozygous or heterozygous Val 129 is related to the phenotype of disease.

We tested 21 unrelated Japanese patients with sporadic CJD, who lack mutations at PrP codons 102, 117, 178, 198 and 200 or insertion mutations seen in inherited spongiform encephalopathies. One is homozygous for Val 129, four are heterozygous, and the remaining sixteen are homozygous for Met 129. Because population data from 179 unrelated Japanese people show a much higher prevalence of codon 129 homozygosity (164 Met/Met 129, 15 Met/Val and no Val/Val), we conclude that Japanese patients with sporadic CJD do not carry an excess of homozygous Val 129 or Met 129 ($\chi^2=1.40$, $P>0.1$). The Japanese population is in Hardy–Weinberg equilibrium for the alleles ($\chi^2=0.34$, $P>0.1$), but the allele frequencies (Met:Val 129, 0.958:0.042) vastly differ from those found in the United Kingdom (0.625:0.375).

It is interesting that clinical and pathological features depend on Val 129. Japanese patients without Val 129 had a shorter clinical course (mean 17.0 months, s.d. 8) with visual impairment, personality change or hallucination at the onset of the disease. Myoclonus and periodic synchronous discharges in the EEG were always observed, but pathology showed no amyloid plaques in the brain using anti-human PrP antiserum. In contrast, all five patients with Val 129 had a longer course of disease (55.6 months, s.d. 30), starting with ataxia, amnesia or parkinsonian symptoms. Not all of them showed myoclonus or periodic synchronous discharges, but all had marked amyloid plaques in the brain which stained with anti-human PrP antiserum. These clinical and pathological features are rather similar to those of Gerstmann–Sträussler–Scheinker disease (GSS). We also tested two unrelated families (N.N. and K.M.) with GSS. One individual from N.N. and two sisters from K.M. had similar clinical and pathological features and were heterozygous for Val 129. Sequence analysis in all CJD and GSS patients with Val 129 revealed no other mutations in PrP. These results suggest that among Japanese CJD patients, individuals who are either homozygous or heterozygous for Val 129 will express the GSS phenotype and should be categorized as GSS, just like patients with sporadic

CJD who have the codon 102 Leu variant of PrP (ref. 2).

Our results differ from those of Palmer *et al.*¹. This discrepancy suggests that other genetic differences between Japanese and UK populations might also be influencing CJD expression and phenotype.

K. DOH-URA

LPVD, Rocky Mountain Laboratories,
NIAID, Hamilton,
Montana 59840, USA

T. KITAMOTO

Y. SAKAKI

J. TATEISHI

Department of Neuropathology,
Neurological Institute and Research
Laboratory for Genetic Information,
Kyushu University, Fukuoka 812, Japan

COLLINGE AND PALMER REPLY — Clearly, PrP Val 129 is rare in the normal Japanese population (15/343 chromosomes); we have found similar gene frequencies among Thais. Therefore it would be very difficult for these workers to replicate our study¹ in their population. Nevertheless, it is of interest that four of the cases described above by Tateishi's group are heterozygotes and that these are 'atypical' CJD cases of long duration (5–10% of CJD patients present like this³). We reported a single heterozygote in the CJD group (which turned out to be a familial case) and four heterozygotes from our suspected CJD group out of a total of 45 cases¹. We have since screened further clinically suspected cases from the UK CJD surveillance study (kindly provided by R. G. Will) and found a further four heterozygotes. We had genotyped these cases without knowing their clinical details. Our prediction, following from our paper¹, was that these heterozygotes would turn out either to be misdiagnosed, unsuspected familial cases with pathogenic prion protein mutations, or 'atypical' CJD of long duration (phenotypically like familial CJD or GSS). This is essentially what we found (two were not CJD, one was confirmed CJD of short duration but with an atypical EEG, and the other five were clinically 'atypical' CJD). Unfortunately, Tateishi and colleagues do not mention the duration of illness in the single Val 129 homozygote, but give an overall mean with heterozygotes.

Both our data and those of Tateishi's group therefore support the idea that classical sporadic, short-duration CJD occurs in homozygotes, with the occasional heterozygotes with the sporadic illness accounting for the 5–10% of atypical cases. This is consistent with the emerging view that interaction between homologous PrP molecules may underlie the disease process and that heterozygotes are partially 'protec-

ted' from the disease.

In addition, all the autosomal dominant inherited prion disease cases are heterozygotes with respect to a pathogenic mutation (at codons 102, 117 and so on), and so also have heterologous PrP molecules. This may also interfere with the efficiency of interaction between PrP molecules and could account for the 'atypical' clinical presentation of the inherited cases (usually described as GSS, a relatively protracted illness in comparison to the dramatic sub-acute presentation of CJD). Interestingly however, codon 200 disease presents as CJD.

Perhaps Tateishi's group has heterozygotes in his CJD patients, despite the

low population frequency, as a result of selection, in that such rare interesting cases are sent to centres with special expertise in these conditions. Finally, with respect to the Indiana kindred of GSS, a missense mutation (at codon 198) has now been reported in this family⁴.

JOHN COLLINGE

MARK PALMER

Department of Biochemistry and
Molecular Genetics,
St Mary's Hospital Medical School,
Norfolk Place, London W2 1PG, UK

1. Palmer, M. S., Dryden, A. J., Hughes, J. T. & Collinge, J. *Nature* **352**, 340–342 (1991).
2. Doh-ura, K., Tateishi, J., Kitamoto, T., Sasaki, H. & Sakaki, Y. *Ann Neurol.* **27**, 121–126 (1990).
3. Brown, P. *et al.* *Ann. Neurol.* **16**, 295–304 (1984).
4. Prusiner, S. B. *Science* **252**, 1515–1522 (1991).

Isodisomy in BWS chromosomes

SIR — Henry *et al.*¹, in demonstrating paternal disomy in cases of Beckwith-Wiedemann syndrome (BWS), have made an important advance in understanding this disease. But the authors do not discuss a surprising aspect of their data: excess homozygosity for chromosome band 11p15.5 markers is detected for the INS and IGF2 loci, but not for the flanking HRAS1 and HBB loci in BWS patients (Table 2 of ref. 1). This observation is difficult to explain if disomy for the entire paternal chromosome has occurred: as the INS/IGF2 region is also a candidate location for the BWS mutation², this raises the question of whether the uniparental disomy is restricted to the 11p15.5 region and has arisen by a localized mechanism.

In uniparental disomy involving an entire chromosome, the mechanism is defined by whether the two centromeres arise from the same parental chromosome (isodisomy) or different chromosomes (heterodisomy). The chance of isodisomy at any particular position along the chromosome depends both on the mechanism of the disomy (meiosis I defects cause centromeric heterodisomy, meiosis II defects centromeric isodisomy), and on the genetic distance of the locus from the centromere (a single intervening recombination tends to reverse the pattern)³. Although this may result in graded changes in observed homozygosity along a population of uniparentally disomic chromosomes, sharp local variations such as that observed between HRAS1 and HBB are not expected (the recombination fraction between these loci is only 0.18 in males⁴). What other mechanism(s) could account for the apparently localized homozygosity in BWS patients?

First, and perhaps most likely, it could be an artefact. Although Henry *et al.*¹ measured the frequency of homozygosity for the INS and IGF2 polymorphisms in a control group, they used published

data⁵ to estimate homozygosities for the flanking HRAS1 and HBB polymorphisms. These estimated homozygosities may be too high (for instance, the published data⁵ suggest a figure for the HBBP/HincII system of 0.36–0.40, lower than the value of 0.51 cited in ref. 1), and it is not clear whether comparable racial groups were sampled. It is important to measure homozygosities directly in controls and to study more patients. In addition, cases with proven isodisomy for 11p15.5 should be tested with polymorphic loci on the long arm of chromosome 11: paternal disomy for such distant loci would imply that it involved the entire chromosome (Angelman syndrome provides an example of the application of this approach⁶).

Supposing that the isodisomy were indeed localized, this would imply a distinct mechanism, presumably involving the copying between chromosomes of a small portion of DNA in band 11p15.5, and occurring either during meiosis or mitosis. It is conceivable that the segment of chromosome copied could carry a pre-existing recessive mutation, explaining the requirement for iso- rather than heterodisomy. Such a mechanism would be of great biological interest and would suggest novel strategies for the localization of the BWS gene(s).

ANDREW O. M. WILKIE

SUSAN MALCOLM

MARCUS E. PEMBREY

Mothercare Unit of Clinical Genetics
and Fetal Medicine,
Institute of Child Health,
London WC1N 1EH, UK

1. Henry, I. *et al.* *Nature* **351**, 665–667 (1991).
2. Ferguson-Smith, A.C., Cattanaach, B.M., Barton, S.C., Beechey, C.V. & Surani, M.A. *Nature* **351**, 667–670 (1991).
3. Engel, E. *Am. J. med. Genet.* **6**, 137–143 (1980).
4. Keats, B., Ott, J. & Conneally, M. *Cytogenet. Cell Genet.* **51**, 459–502 (1989).
5. Kidd, K.K. *et al.* *Cytogenet. Cell Genet.* **51**, 622–947 (1989).
6. Malcolm, S. *et al.* *Lancet* **337**, 694–697 (1991).

Usage and abuse

Dorothy Nelkin

Value-Free Science? Purity and Power in Modern Knowledge. By Robert N. Proctor. Harvard University Press: 1991. Pp. 352. \$34.95, £27.95.

How often have we heard this refrain? It is not science that is the source of problems but only its abuse; science, in itself, is neutral if uncontaminated with values; science can be used for good or for evil, to cure or to kill, to develop or to destroy. In *Value-Free Science?*, Robert Proctor takes on the ambitious task of tracing the history of the idea of neutrality in science, explaining its origins and its role. He argues that neutrality has been the bargain struck by scientists to pay for social legitimation in the eyes of the Church and, later, of the state. Value-free science, in effect, has been the price of autonomy, serving to shield science from outside control.

He documents this theme starting with the platonic separation of theory and practice. He describes how the practical vision of knowledge, expressed in the philosophy of Francis Bacon, required a negotiation that would permit freedom of inquiry as scientists participated in public affairs. He explores the subsequent separation of fact and value, reviewing weberian concerns about academic freedom, censorship and the inclination to pass off political opinion as fact. His most powerful analysis emerges from his work on Nazi Germany, demonstrating the implications of using neutrality as a way to escape explosive political questions.

Proctor believes that claims of neutrality arise as a response to broad changes in the social and economic context of science and, in particular, the challenges associated with the rise of industrial, military and state support. When social movements such as marxism and feminism questioned the autonomy of science, claims of neutrality became a protection. In effect, value neutrality has served as an ideology of 'science under siege', allowing scientists to maintain autonomy in response to pressure by governments for censorship, by industry for practical results, and by social movements for relevance. The exclusion of values, of morality from science, acts as a shield, a way to avoid external intervention in the affairs of science. But, more than a shield, says Proctor, this ideology is also a weapon that has been used against feminists, social darwinists, socialists and others who have tried to politicize science. Just as the natural philosophers maintained autonomy by avoiding matters of morals and religion, so modern scientists have tried to define themselves

as outside the political and moral arena.

Proctor sees today's perpetuation of the ideal of neutrality as a way to blunt social criticism, to maintain the existing structure of power. Essential to his argument is the difference between neutrality (whether science takes a stand) and objectivity (whether science merits claims to reliability). By conflating objectivity with neutrality, scientists have maintained an ambiguous position, avoiding political commitment while engaging themselves in practical affairs: "Science . . . serves two masters; the cause of truth, but also the practical needs of those who pay." Proctor's concern lies in the hypocrisy of this position; although claiming neutrality, science is in fact serving special interests and reinforcing social patterns of dominance and exclusion.

Part of Proctor's mission is to influence the sociology of science. He applauds the recent efforts to examine the social construction of scientific ideas; "science is not a self-sustaining filiation of disembodied ideas but a social process with material and social prerequisites." But he would go well beyond this approach, for he sees fundamental problems in the very structure and practice of

science. Thus, he calls for a "moralistic science theory", a science based on advocacy and commitment. This, he claims, would not compromise objectivity. It simply means that science must "pay attention to the concrete forms of suffering and injustice in the world". Challenging the legacy of "detached indifference" as a political ideology, a way to camouflage interests and escape from commitments, he seeks to base science on "politicized moralism". For in any case, there is much evidence of the politicization of science, for example, in its engagement with military developments, hereditary studies and agricultural research.

Unfortunately, Proctor's analysis stops at this point. After reviewing an enormous amount of material on the history of neutrality, synthesizing a vast and often diffuse literature on the relationship of science to power, and offering a powerful critique of the ideological dimensions of scientific claims, Proctor stops. He neither provides us with a "moralistic science theory" nor tries to develop its agenda. What would science look like if it were to evolve as he would have it? How would it be organized and supported? What problems should it address? Without working out either a theory or an agenda, this book becomes more of a moral outcry than a political analysis. Yet it raises provocative questions for others to explore. □

Dorothy Nelkin is in the Department of Sociology, New York University, 269 Mercer Street, New York, New York 10003, USA.

An Arctic route to the East



The quest for the Northwest Passage was one of the greatest navigational challenges for early mariners. For more than 500 years explorers tried in vain to find a commercial route to the markets of the Orient and it was not until 1906 that Roald Amundsen successfully navigated the passage. This illustration shows the British explorers John Ross and William Edward Parry encountering Greenland Inuit in 1818, and the contrast in clothing indicates how badly prepared the officers were. *Northwest Passage* by E. Stuzik and M. Beedell beautifully chronicles these early ventures. Published by Cassell, price £16.95. □

Sisters in arms

Jon Seger

The Social Biology of Wasps. Edited by Kenneth G. Ross and Robert W. Matthews. *Cornell University Press: 1991. Pp. 678. \$72.50 (hbk), \$34.95 (pbk).*

SOCIAL wasps seem more familiar — somehow easier to understand and to identify with — than either social bees (which one seldom notices anyway) or ants (which might as well have come from Mars). Why should this be so? In a technical sense all three groups are 'wasps', that is, members of the suborder Aculeata of the insect order Hymenoptera, and thus closely related to each other and equally distant from us. They all live as multigenerational families in permanent nests, where most females act as sterile helpers, working on behalf of one or a few reproductive colony members (usually their mothers, aunts or sisters). Discussions of insect sociality tend to emphasize these similarities as well as the search for correspondingly general explanations of helping behaviour and all that it entails. Such generalities exist and are of great interest, but as our knowledge of social insects becomes both broader and deeper, the differences between the main groups seem more and more profound. One unstated message of this splendid book is that the social wasps may be, in some respects, more like six-legged, four-winged hominids than they are like other social insects. Intelligent, adaptable ecological generalists and skilled makers of artifice, they are, above all, political animals (*Polistes* means just that) for whom the tension between cooperation and self-aggrandizement is the overriding condition of daily life.

The book contains 17 well written, carefully edited, lovingly illustrated chapters by leading students of wasp evolution and behaviour, 1,500 references and a masterful introduction by Mary Jane West Eberhard that places the current enviable state of wasp sociobiology in the context of its long and distinguished history. The chapters fall into two sections of roughly equal length; the first is organized systematically and the second thematically. Many of the chapters in each section are the best available reviews of their kind. Some are almost the only reviews of their kind, for example, Carpenter's opening chapter on the phylogeny of Vespidae and the evolution of social behaviour in the family (defined broadly to include all the vespoid social wasps and their solitary allies) and Wenzel's

chapter on nest architecture.

The systematic section includes chapters on all the main groups of vespids, from the solitary and presocial subfamilies (D. P. Cowan) to the most advanced Vespinae (A. Greene), and includes welcome treatments of under-studied and under-publicized groups such as Stenogastrinae (S. Turillazzi) and *Ropalidia* (R. Gadagkar). Each of these chapters covers various aspects of biogeography, natural history and behaviour, including nest establishment and architecture, colony cycles, dominance relations, brood rearing, sex ratios and enemies. The special-topics section then slices the cake the other way, with chapters focusing on some of the recurrent themes of the previous section, such as competition during and after colony establishment (P.-F. Röseler, C. K. Starr), regulation of queen number (J. P. Spradbery), behavioural specialization among colony members (R. L. Jeanne) and the many functions of exocrine glands (H. A. Downing). The final chapter (R. W. Matthews) is a review of social behaviour in the Sphecidae, a very diverse group of mostly (but not entirely) solitary wasps that is more closely related to the bees than to the Vespidae.

One might expect a tour of taxa to be purely descriptive and dull, and a festival of themes to be arid. In fact, the systematic chapters are well informed by theory and on some points do a better job of explaining it than do the thematic chapters, which are, to their credit, rich in biological detail. The lack of a chapter on kin selection and sex ratios could be viewed as one of the book's few weaknesses, but the extent to which this material is incorporated organically into many chapters shows that the field is reaching a new level of sophistication and maturity, in which there is less and less segregation of big questions from small facts. Many topics come up repeatedly in different contexts, but the editors are right to make no apology for this; a field can be displayed in a different way in an edited volume than it typically is in a textbook, and this volume shows that the difference can be a strength. Anyone with even a casual interest in social insects should read it. □

Jon Seger is in the Department of Biology, University of Utah, Salt Lake City, Utah 84112, USA.

■ Two books on bees have recently been published: *The Behaviour and Physiology of Bees* edited by L. J. Goodman and R. C. Fisher and based on a 1990 colloquium organized by the Royal Entomological Society and International Bee Research Association (C.A.B. International, price £49.50, \$94); and *Bees of the World* by C. O'Toole and A. Raw, aimed at a more general audience. (Cassell, price £16.95.)

Considering cosmology

David W. Hughes

Our Universes. By Denys Wilkinson. *Columbia University Press: 1991. Pp. 213. \$52.*

THE gap between modern philosophy and natural philosophy seems to be ever-widening. Many of those with the degrees of DPhil and PhD spend time apologizing for lack of knowledge of philosophy, explaining that they are mere scientists. In my speciality — astrophysics — the most philosophical of the disciplines is cosmology. Rude astrophysicists suggest that the philosophical content of a discipline is directly proportional to the sparsity of facts. Cosmology wins hands-down in this department.

Nonetheless, there are three good facts. First, the Universe is expanding, a fact discovered by Edwin Powell Hubble in 1929. Second, 25 per cent of the mass of the Universe is helium, most of the rest being hydrogen; the nucleosynthetic significance of this was recognized by William Fowler and Fred Hoyle in the mid-1950s. Third, the Universe is permeated by a three-degree microwave background radiation, discovered by Arno Penzias and Robert Wilson in 1965. Together these facts point to the co-creation of matter and space in a hot Big Bang about 15,000 million years ago.

But there is a long list of problems. First, the Universe is extremely bright, there being 10^9 photons per nucleon. Even in a supernova explosion, the ratio of photons to nucleons is 100 times smaller. Second, the universal space is very nearly flat, and so its expansion energy and gravitational energy are nearly equal and its density is also close to the closure density (the density that will make the Universe just stop expanding when it has reached an infinite size). Third, the background radiation is extraordinarily uniform (better than 0.01 per cent), and the reason why galaxies were formed is thus a bit of a mystery. Fourth, coupled with the last problem is the need to explain why matter has a lumpy distribution but not as lumpy as one would expect if it initially had a distribution equivalent to thermodynamic equilibrium. Fifth, the expectation of cosmic democracy has been thwarted by the complete imbalance between matter and antimatter. And sixth, we seem to have lost the magnetic monopoles and the cosmological constant.

These are the facts and conundrums that give rise to the philosophical musings of Sir Denys Wilkinson in this book. Sir Denys recently retired from being the

vice-chancellor of the University of Sussex and a professor of nuclear and experimental physics at Oxford University. In 1989 he delivered the George B. Pegram Lecture at the Brookhaven National Laboratory; *Our Universes* is the resulting book.

Wilkinson's view of cosmology seems to be greatly coloured by the writings of Bishop Berkeley. Much is made of the relationship between the physical 'real' Universe (if such a thing exists) and the sentient mind that both looks at it and thinks about it. The author also stresses the importance of the nuclear processes that took place in the first fraction of a second after the Big Bang, and the fact that these could have gone many different ways and produced many different end-products as well as many different disconnected universes at the same time.

Much is made of the inconsistency between the copernican principle, which implies that we do not occupy a privileged position and that everything must perforce be commonplace, and the anthropic principle, which must have seriously tied the hands of the creator of the Universe if He wished it to be occupied by living sentient beings. Wilkinson stresses just how lucky we are to be here today. For example, if the charges on the electron and proton were a factor of three higher, the periodic table would end at the element boron; if the charge on the electron exceeded that on the proton by one part in 10^{18} , the overall electrical repulsion between a human and Earth would equal the gravitational attraction and we would float off into space; if there had been any other than three spatial dimensions, stellar, planetary and atomic orbits would be unstable; if the proton were heavier than the neutron the Universe would consist entirely of helium; if the ratio between the density of the Universe and the closure density were greater than ten then the Universe would have collapsed before stars could have been formed; and if that ratio had been less than 10^{-3} , the Universe would have expanded so rapidly that matter could not have condensed into galaxies.

This is a delightful book in which the pace of the story-telling never flags. It is also most refreshing to see cosmology through the eyes of a nuclear physicist. It is clear from the text that Wilkinson is not only an excellent lecturer but also a profound thinker. He has the gift of encouraging his reader to think too. I found myself continually making mental notes to read more about some of the topics. Here the lack of a reference list will be a distinct drawback. □

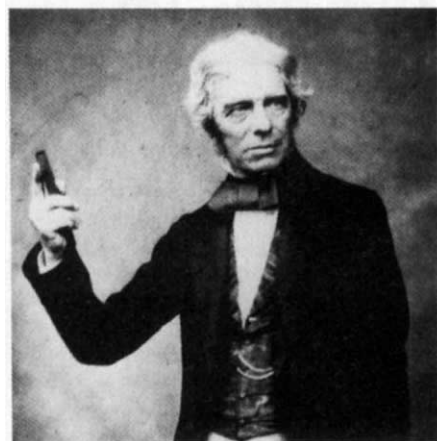
David W. Hughes is in the Department of Physics, University of Sheffield, Sheffield S3 7RH, UK.

From candles to chemistry

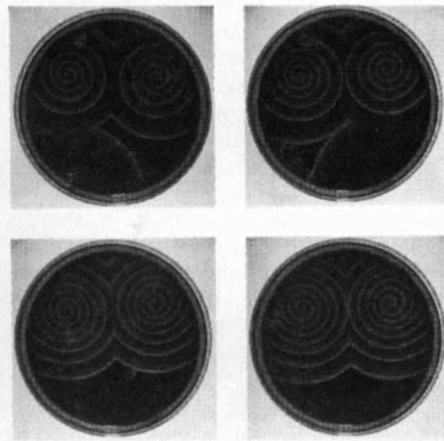
J. P. Simons

Atoms, Electrons and Change. By P. W. Atkins: W. H. Freeman: 1991. Pp. 243 £16.95, \$32.95.

ONE hundred and fifty years ago Michael Faraday presented a series of discourses on 'the chemical history of a candle', that most symbolic and evocative of human artefacts. This year, when so many are celebrating the bicentenary of his birth — at Westminster Abbey; at the crucible of his genius, the Royal



most affectingly when it is borne on the wave of excitement and romance that drives the best research scientists and provides them with their inspiration. Reading Atkins' book is rather like riding a surfboard on that wave. It is an exhilarating ride. The book is very glossy but beautifully produced, with many



Prime movers — Michael Faraday (1791–1867) and the Belousov – Zhabotinskii oscillating reaction.

Institution; at the Institution of Electrical Engineers; and even at the Bank of England — Peter Atkins has courageously tackled again the task that Faraday set himself. That task is to see the complexity (and underlying simplicity) of chemical change through the light of a candle's flame but with the insights of the present century. And, of course, the insights are many. Chemical thermodynamics was developed in the 1870s by Willard Gibbs, a few years after Faraday's death. The notions of electronic and molecular structure and their understanding in terms of quantum mechanics arrived over the next half a century. The primary steps in photochemical change, not least that most primary of processes, photosynthesis, were incomprehensible until the advent of quantum theory. Even Maxwell's demons, which can stop a molecule in its tracks, have now materialized with the advent of the laser: femtosecond lasers can provide real-time snapshots of molecules falling apart.

The ability to understand large-scale chemical transformations (and thereby to control them) in terms of the microscopic motions of atoms, molecules and their constituent electrons is one of the greatest achievements of this century. Indeed, it is the basic foundation of much of our physical and economic well-being. This cannot be said too loudly or too often, but the message comes over

pictorial aids to understanding, and contains illustrations that are like the 'stills' advertising a series of 'demonstration lectures', which is what the nine free-standing chapters really are. (Faraday only needed six, but a lot has happened since the 1850s.)

The book opens with an introductory essay on what Faraday did and did not know about chemical change, in particular the transformations in a candle flame that cycle carbon from its reduced to its elemental and oxidized states — from hydrocarbon wax or fats, to soot, carbon dioxide and thence to carbonic acid and limestone or back to carbohydrates by way of photosynthesis. This cycle forms the warp upon which much of the book's story is woven. It begins where Faraday stopped, with a layman's guide to the electronic structure of atoms; the notions of electron orbitals, spin and Pauli's exclusion principle; the nature of chemical bonding and the crucial importance of phase (designated as 'colour'). The description of bonding using the linear combination of atomic orbitals is illustrated using carbon dioxide and the carbonate ion as examples.

In the third chapter, Atkins invites readers to contemplate their aqueous chemistry and uses it to trace the historical development of the concept of acids and bases, concluding with the simple idea that they correspond to (positive)

holes and electrons. On the way, the reader also encounters redox and free-radical reactions, which Atkins links together in a dazzling virtuoso display — a little too dazzling perhaps when the final conclusion is reached: “all the phenomena of the world are merely the reactions of one kind of acid with one kind of base”.

The driving force for chemical change leads us into thermodynamics, Clausius's principle, and the tendency towards chaos, and the virtues of coupled reactions. Although the fourth chapter opens with a discussion of Gibbs, his free energy is never encountered. Chemical equilibria are soon followed by the idea of reaction rates and energy barriers to reactions. The story moves swiftly from experimental observation to theoretical understanding in terms of molecular collisions, the dynamics involved in the making and breaking of chemical bonds, classical trajectories, Polanyi's rules for energy use and disposal (although Polanyi is not mentioned), tunnelling, isotope effects, diffusion-controlled reactions, catalysis at surfaces and ‘femtochemistry’. All this is accomplished in 25 sparkling pages — a *tour de force* and a delight.

Atkins begins the sixth chapter with an introduction to chemical reaction kinetics and rate equations, but halfway through the going becomes decidedly tough as autocatalysis, nonlinear rate processes, oscillating reactions and chaos are encountered. The medium is the famous but highly complex Belousov-Zhabotinskii reaction which Atkins discusses in some detail. Light relief comes at the end of the chapter with the association of concentration waves and animal-colouring waves that produce, for example, tigers' stripes and leopards' spots. This is a giant leap from Harcourt and Esson's rate law.

The next two chapters deal with the mechanisms of organic chemical reactions, and their description in terms of reactive intermediates, changes in electron distributions, proton transfers and sequences of small changes that can be amplified to larger ones. Atkins clearly explains the profound importance of orbital symmetry in providing a rationale for the preferred reactivity and stereochemistry of concerted molecular reactions where ring structures are generated, opened or closed. But without the Woodward-Hoffmann rules, which Atkins does not detail, it is not easy to explain why ethylene molecules do not readily bond together or why the addition of ethylene to butadiene helped to make Otto Diels and Kurt Alder famous.

The book closes with a chapter on photobiology, embracing luminescence, lasers and photosynthesis (the sections

on lasers might have been better placed with those on reaction dynamics). The account of photosynthesis, although pulling no punches, is clearly developed and illustrated, and forms a fitting conclusion to a grand concept. *Atoms, Electrons and Change* deserves to be read widely, even if the author believes its conception to be only “the consequence of a chemical reaction”. □

J. P. Simons is in the Department of Chemistry, University of Nottingham, University Park, Nottingham NG7 2RD, UK.

Touching on the future

Claes G. H. Rooth

Global Changes of the Past. Edited by Ray Bradley. *Office for Interdisciplinary Earth Studies**. 1991. Pp. 514. \$38.

Paleoclimatology. By Thomas J. Crowley and Gerald R. North. *Oxford University Press*: 1991. Pp. 346. £45, \$59.95.

THE study of the character and causes of climate variations, as well as efforts to divine their future trends, has blossomed over the past few decades from an obscure activity pursued by a few enthusiasts into a rapidly growing and exciting area of science. The foundations for a physically and dynamically based theory of climate were laid in the recent maturation of several quite disparate specialities such as micropalaeontology, biogeochemistry and weather forecasting using computers.

Quantification of the different features of climate relies on statistical descriptions of the highly variable weather scene. Although the definition of these features in legal and economic practice is usually based on a 30-year period, we are increasingly aware of their long-term variability. As guides to the facts about such variability, these two books should be useful to a wide variety of users, be they researchers, policy-makers or broadly interested seekers of information about our environment. What awaits us in the future is touched on only in the broadest terms, however, and appropriately so, because natural variability, the causes of which are still poorly understood, is difficult to distinguish from the effects on the environment for which we humans have been responsible.

Our knowledge of past climates is still largely based on ingenious use of circumstantial evidence, but the interpretation

of this evidence is increasingly restricted by what is allowed by physical models. But although there are some absolute constraints, the models remain highly abstracted and so baconian interplay between developments in observation and theory is called for. This is brought forth in *Global Changes of the Past*, a collection of 20 tutorial presentations prepared for the second Global Change Institute meeting held at Boulder, Colorado, in 1992.

The volume also includes a set of working-group assessments of future climate, as well as recommendations about what action we should be taking. These are not very different from much that has been expressed elsewhere, but they provide a compact overview of the field, and thus a useful introduction for the novice. Although there is no overall consistency in style, the tutorial chapters contain state-of-the-art reviews of most if not all of the key components of current climate research. In general they are well written, illustrated and referenced. The book's main limitations lie in the areas of computational climate modelling and research into biogeochemical processes.

In *Paleoclimatology*, Crowley and North make considerably heavier fare, and succeed admirably in presenting an authoritative and lucid review of diverse material. The complementary backgrounds of the authors in geology and physics have enabled them to sweeten the usually dry flavour of reviews of geological and micropalaeontological data with plenty of dashes of good ideas and insights. For this the reader is prepared by an introductory review of elementary models of climate systems, and of the basic dynamics (including numerical modelling) of the large-scale oceanic and atmospheric circulations. At the same time, the otherwise mostly speculative forecasting by computer models of possible climates for epochs far from our own is given credence by being presented in a well-integrated way. The book is illustrated generously and well, which with the authors' easily flowing writing style makes it rewarding to the browser and keeps it interesting for those searching for a deeper understanding. But because of its broad scope, the serious student will probably need to search further. In this respect, the book's listing of more than 1,200 text references and almost 40 reference works will come in handy. Nevertheless, the work is complete enough to be both a standard reference for an audience ranging from researchers and students to environmental planners, and an excellent source for science communicators. □

Claes G. H. Rooth is at the School of Marine and Atmospheric Science, University of Miami, Miami, Florida 33149, USA.

* University Corporation for Atmospheric Research, PO Box 3000, Boulder, Colorado 80307 3000, USA.

Gamma rays and energetic particles from primordial black holes

F. Halzen, E. Zas, J. H. MacGibbon & T. C. Weekes

Black holes of almost arbitrarily small mass may have formed in the very early Universe. Their presence today would be revealed by the energetic radiation they would produce by means of the quantum-gravitational Hawking mechanism, allowing observational limits to be set on their density today and on their past significance.

UNDER present cosmic conditions, black holes can form only by self-gravitational collapse, and since masses of about three solar masses ($M_{\odot}$) or less are stable when cold in the form of neutron stars or white dwarfs, modern black holes must exceed a few $M_{\odot}$ in mass. But in the early Universe primordial black holes (PBHs) of almost arbitrarily small mass may have formed. The simplest mechanism^{1,2} is the collapse of overdense regions in a universe with significant density fluctuations; the resulting PBHs have masses roughly equal to the mass contained within the cosmic horizon (essentially c times the age of the Universe), which can be as low as the Planck mass $(\hbar c/G)^{1/2} \approx 5 \times 10^{-11}$ g. More exotic formation mechanisms include cosmic phase transitions³ or the collapse of loops of cosmic string^{4,5}. The presence of any such PBHs in the Universe today would go largely unnoticed were it not for the Hawking radiation from black holes⁶, a phenomenon arising out of the union of general relativity and quantum mechanics, according to which black holes can radiate particles whose Compton wavelength is less than the Schwarzschild radius of the black hole (see Box 1).

For black holes of solar-mass size, Hawking radiation is quite negligible, but for small enough PBHs, it becomes the controlling influence in the black hole's evolution. PBHs of $\sim 5 \times 10^{14}$ g or less will indeed have evaporated entirely in the 10^{10} years or so of the Universe's history. PBHs a little more massive than this initially will still be evaporating, at a rate large enough that the stream of energetic particles and radiation they emit can be turned into a strong observational limit on their presence. The cosmic ray and γ -ray backgrounds place an upper limit on the integrated density of PBHs with initial masses in this range of $\sim 10^{-8} \rho_c$, where ρ_c is the critical density required to close the Universe, but even at this level their emissions could make a significant contribution to the observed extragalactic γ -ray and interstellar positron, electron and antiproton backgrounds. PBHs whose initial mass exceeded 5×10^{14} g by two orders of magnitude or more can provide ρ_c but escape detection because their emission would be concealed by the larger observed backgrounds at energies of 100 keV or less.

Moreover, a population of PBHs whose influence is small today may have been more important earlier in cosmic history. Too much radiation from PBHs could perturb the usual picture of cosmological nucleosynthesis⁷, distort the microwave background⁸ and produce too much entropy in relation to the matter density of the Universe (for a recent review, see ref. 9). In general, limits on the density of PBHs, now or at earlier times, can be used to provide information on the homogeneity and isotropy of the very early Universe, when they were formed¹⁰.

Searches for PBHs attempt either to detect a diffuse photon background from a distribution of PBHs or to search directly for the final emission stage of individual holes. Here we review the status of the search for PBHs and discuss strategies for improving the existing bounds on black-hole abundance, obtained from diffuse megaelectronvolt γ -ray data^{11,12}, by exploiting air-shower telescopes to search for the final γ -ray emission. Particle emission by PBHs was initially researched

over a decade ago^{6,10,11} and has been recently reconsidered by MacGibbon, Carr and Webber¹²⁻¹⁴ in light of the discovery of the quark-lepton structure of matter. They assume that asymptotically free quarks and gluons are emitted which hadronize into jets of astrophysically stable photons, neutrinos, electrons, positrons, protons and antiprotons. The jet emission mimics particle production by an electron-positron collider. The ultimate behaviour of a black hole at very high temperatures is particularly interesting, as it is dictated by particle physics beyond the range of current particle accelerators. We will calculate the experimental signatures of Hawking radiation in the

BOX 1 Hawking radiation from black holes

The radiation of particles by black holes can be qualitatively understood by analogy with the production of e^+e^- pairs in the presence of strong electric fields. In quantum mechanics, virtual e^+e^- are continuously created and destroyed in the vacuum. A strong field can separate the charges, and as a result, particles can tunnel through the quantum barrier and pop out of the vacuum as real particles. The threshold field strength is given by

$$eE\lambda > 2m_e c^2 \quad (1)$$

The left-hand side is the work done in separating the particles by a Compton wavelength λ . The right-hand side is the energy required to create the particles. The energy of the particles can be estimated as follows

$$E = kT \propto \rho c \propto \frac{\hbar}{\lambda} c \propto \frac{\hbar c^3}{2G} M^{-1} \quad (2)$$

where we took the horizon or Schwarzschild radius of the black hole $2GM/c^2$ as the relevant Compton wavelength. The luminosity is determined by the area multiplied by the energy

$$L \equiv \frac{dM}{dt} = (4\pi\lambda^2)(aT^4) \propto M^{-2} \quad (3)$$

The energy is determined by the Stefan-Boltzman relation and the proportionality constant, a , in (3) therefore contains the number of degrees of freedom of the particle emitted. The timescale associated with this radiation is

$$t = \frac{M}{dM/dt} \propto M^3 \quad (4)$$

These dimensional estimates can be summarized as follows

$$T \approx 100 \text{ MeV} \left(\frac{10^{15} \text{ g}}{M} \right) \quad (5)$$

$$L = 10^{20} \text{ erg s}^{-1} \left(\frac{10^{15} \text{ g}}{M} \right)^{-2} \quad (6)$$

$$t \approx 10^{10} \text{ yr} \left(\frac{M}{10^{15} \text{ g}} \right)^3 \quad (7)$$

Small black holes with mass of $\leq 10^{15}$ g or less should be observable. They are characterized by high temperatures and luminosities. They moreover have lifetimes such that they evaporate, possibly explosively, within the lifetime of the universe.

context of the standard model of quarks and leptons and also illustrate the effect of particle degrees of freedom beyond the standard model. The standard model is known to be incomplete, and the additional degrees of freedom are a subject of intense speculation. Searching for them constitutes the primary mission of future accelerators such as the SSC.

On the experimental side, a new generation of particle detectors, covering the TeV (10^{12} eV) and PeV (10^{15} eV) energy ranges, and new γ -ray telescopes such as the EGRET detector on the Gamma Ray Observatory greatly increase the potential sensitivity of searches for Hawking radiation (Box 2). The astronomical activity of such instruments has almost exclusively been focused on the search for point sources of ultra-energetic particles¹⁵⁻¹⁸. These detectors can probe diffuse high-energy photon backgrounds produced by galactic and extragalactic cosmic rays^{19,20}. Here we emphasize another measurement of cosmological relevance, namely the search for Hawking radiation from black holes.

BOX 2 TeV- and PeV-astronomy

The predicted fluxes of TeV and PeV cosmic photons are smaller than the cosmic ray flux in the corresponding energy range. The photons are absorbed by the atmosphere, and detectors of area $\sim 1 \text{ m}^2$ carried by rockets or satellites are not sensitive to the small fluxes discussed here. Fortunately the same atmosphere that shields the photons from Earth constitutes a calorimeter medium which can be exploited to detect them. The atmosphere represents a layer of material with a column density of $1,000 \text{ g cm}^{-2}$, corresponding to ~ 25 radiation lengths. The primary photons initiate showers which dissipate their energy into a vast number of electrons and photons, which are absorbed by the atmosphere at a height of $\sim 10 \text{ km}$. Only the energy dissipated in weakly interacting neutrinos and muons, which are the decay products of pions photoproduced in the cascades, penetrates to observation level on a mountain top or at sea level. Muons lose 2 MeV of their energy through ionization for every 1 g cm^{-2} of matter traversed. To penetrate 1,000 g of atmosphere they must have started with at least 2 GeV energy, and a parent primary must have an energy of at least 10 GeV to be detected via sea-level muons.

The showers initiated by cosmic photons with energies exceeding 100 TeV have a lateral spread of $\sim 0.1\text{--}1 \text{ km}$, because the secondary particles diverge as a result of their transverse momentum and secondary interactions in the atmosphere. An 'extensive air-shower array' of, say, 100 m^2 can detect showers over an effective area of 10 km^2 , a gain in sensitivity of 10^5 , because only a fraction of the shower must traverse the array of detectors to be detected. The trajectory of the primary particle need not be contained inside the array. The array itself can consist of relatively widely separated particle detectors (usually a scintillator viewed by a phototube) sampling the 100-m^2 area. Enough particles survive to mountain (sea-level) altitudes for showers of 10 (100) TeV primary energy to be detected in this way.

Another technique consists of using mirrors viewed by phototubes collecting the Čerenkov light produced by shower particles in the atmosphere. Showers with energies as low as 0.1 TeV can be detected using this technique. A 100-GeV shower still produces several hundred electrons at an altitude of 10 km. A mirror with 2° aperture views showers over an effective area of 10^5 m^2 at an altitude of 10 km on a mountain. Recent experiments have been successful in distinguishing the γ -ray showers from the much more numerous proton showers on the basis of the size of the Čerenkov light image.

At higher energies a popular technique is to measure the electromagnetic and muon component simultaneously in a single shower. The muon detector can be a GeV-muon detector at or near the surface, or a deep underground TeV-muon detector triggered by the air-shower array. The simultaneous determination of the number of electrons and muons in the cascade could lead to the determination of the nature of the primary particle—that is, whether it is a proton or a heavy nucleus. Another possibility to be explored with such a hybrid detector is to do γ -ray astronomy in the presence of cosmic-ray backgrounds by selecting showers poor in muons. Reduced pion production in a photon cascade leads us to expect that the number of muons in a photon shower is less than 5% of the number in a proton shower with similar energy or similar number of electrons³⁶. Muon-poor astronomy therefore requires measurement of the muon and electron number in each individual shower.

Hawking radiation and emission models

Hawking showed that an uncharged, non-rotating black hole emits particles with energy in the range $(E, E + dE)$ at a rate¹⁶

$$\frac{d^2 N}{dt dE} = \frac{\Gamma_s}{2\pi\hbar} \left[\exp\left(\frac{8\pi GME}{\hbar c^3}\right) - (-1)^{2s} \right]^{-1} \quad (1)$$

per state of angular momentum and spin. Here M is the mass of the hole, s is the particle spin, and Γ_s , the absorption coefficient, is in general a function of s , E and M (ref. 21). The absorption coefficient is the probability that the particle would be absorbed if it were incident in this state on the black hole. It appears in the emission formula on account of detailed balance between emission and absorption. In the limit $ME \gg 1$, the instantaneous emission is a function of ME only and the spectrum is that of a black body with a temperature of

$$T \approx 1.06 \times 10^{13} \left[\frac{1 \text{ g}}{M} \right] \text{ GeV} \quad (2)$$

The instantaneous photon flux peaks at $ME \approx 6.12 \times 10^{13} \text{ g GeV}$. Integrated over all energies, the photon flux is

$$dN/dt = 5.97 \times 10^{34} \left[\frac{1 \text{ g}}{M} \right] s^{-1} \quad (3)$$

The instantaneous flux of electrons and positrons peaks at a 50% higher value of ME and the integrated combined $e^\pm$ flux is a factor of 6–7 larger than the photon flux. The detailed spectra, which we will use as input into calculations throughout this review, have been calculated by Page^{21,22} for spin- $\frac{1}{2}$ and spin-1 particles, and by Elster^{23,24} and Simpkins²⁵ for spin-0 particles.

The black hole loses mass at a rate²¹

$$\frac{dM}{dt} = -\frac{\alpha(M)}{M^2} \quad (4)$$

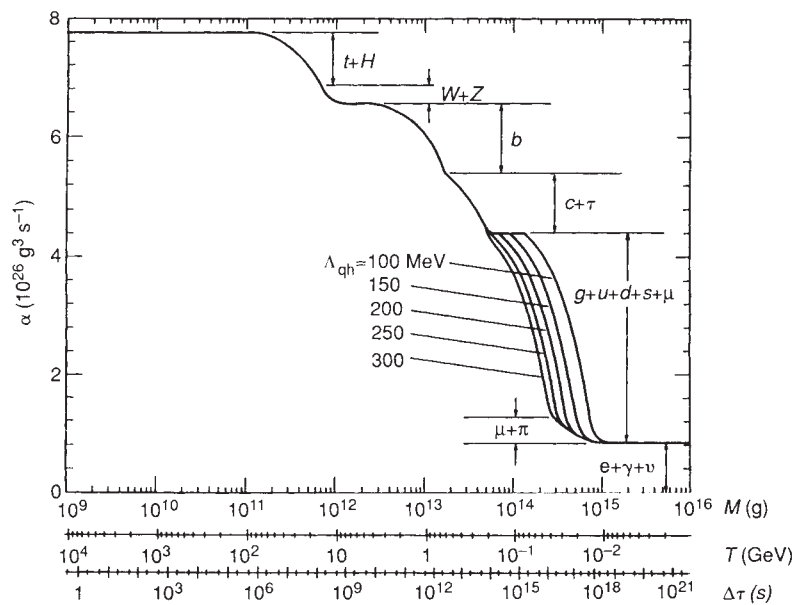
where $\alpha(M)$ accounts for the degrees of freedom of each emitted particle contributing to the energy loss. As the hole radiates, its temperature rises and α increases smoothly at the rest-mass threshold for each new massive particle (Fig. 1). Above each threshold, α is approximately

$$\alpha \approx [7.8 d_{s=1/2} + 3.1 d_{s=1}] \times 10^{24} \text{ g}^3 \text{ s}^{-1} \quad (5)$$

where $d_{s=1/2}$ and $d_{s=1}$ are the number of degrees of freedom (spin, charge and colour) of the emitted particles. For the standard model with three fermionic generations, we have $d_{s=1} = 27$ and $d_{s=1/2} = 90$ at $T \approx 100 \text{ GeV}$, and so $\alpha \approx 7.8 \times 10^{26} \text{ g}^3 \text{ s}^{-1}$ (if we ignore the contributions from the Higgs scalars and gravitons, which have not yet been confirmed experimentally).

Not only does the standard model dictate the high-energy behaviour of the hole, but quantum chromodynamics (QCD) modifies the flux calculations above $\sim 100 \text{ MeV}$. Quarks and gluons, rather than pions and heavier hadrons, should be emitted once the temperature of the black hole (or the peak in the flux distribution) exceeds the quark-gluon deconfinement temperature $\Lambda_{\text{qh}} \approx 100\text{--}300 \text{ MeV}$. Of the hadrons, only pions will be directly emitted below Λ_{qh} . This, of course, only occurs if Λ_{qh} , whose value is not accurately known, exceeds the rest mass of the pion. The emission can be understood by analogy with hadron production in e^+e^- collisions. A virtual photon, produced at rest by e^+ and e^- beams colliding head on, can decay into a quark-antiquark pair. When the pair separate by a distance of $\sim 1 \text{ Fermi}$ (10^{-15} m), the colour interactions between the quarks become exceedingly strong, and violent forces decelerate the quarks. As a result they radiate hadrons (mostly light pions) just as a decelerated charge emits photons by bremsstrahlung. The original quark is never seen; only a 'jet' of pions and other colourless hadrons hits the detector. Similarly a Fermi-size black hole will initially radiate quarks once their energy exceeds Λ_{qh} . The time between emissions is short enough so that we can

FIG. 1 The parameter α counting the degrees of freedom of the Hawking mass radiation as a function of the black-hole temperature and lifetime. Λ_{qh} is the quark-hadron deconfinement scale. The contribution from each particle species is indicated.



neglect interactions between successive emissions before fragmentation into pions occurs^{13,26}. The emitted quarks and gluons can therefore be treated as asymptotically free partons hadronizing to produce jets of pions far beyond the black hole horizon. This is analogous to the production of jets in e^+e^- annihilation. These jets are described by empirical fragmentation functions which represent the way in which the momentum of an initial quark or gluon is distributed among the fragmenting hadrons. These fragmentation functions can be calculated from convenient parametrizations of accelerator data or by Monte Carlo techniques based on QCD¹³.

In this picture it is straightforward to convolve equation (1), describing the emission of quarks and gluons, with the empirical fragmentation functions to obtain the instantaneous emission spectra. In the astrophysical case, the pions will decay into stable $e^\pm$, γ and $\nu\bar{\nu}$. The jets will also produce protons and antiprotons. The black hole also emits direct $e^\pm$, γ and $\nu\bar{\nu}$ fluxes as in equation (1). As an illustration, we show in Fig. 2 the total particle fluxes radiated by a black hole of temperature $T = 10$ GeV. The photon spectrum has a dominant peak around 100 MeV coming from jet pion decays and a smaller bump at an energy of $\sim 5T$ coming from direct γ -ray emission. The β -decay of jet neutrons produces low-energy bumps in the lepton spectra around 1 MeV.

The value of α in equation (5) represents a minimum bound at energies exceeding those probed by particle accelerators. For example, if supersymmetry is the underlying theory for high-energy elementary particles, α should increase by at least a factor of three. There have been many speculations on physics beyond the standard model. We will resist the temptation of listing their predictions and instead choose an extreme example^{27,28} where the number of degrees of freedom grows exponentially with mass, m

$$N \approx m^{-5/2} \exp \left[\frac{m}{\Lambda} \right] \quad (6)$$

The Hagedorn picture, which is motivated by the apparent exponential increase in hadronic resonances, is an example of such a theory. Although not favoured, it is not yet ruled out by particle physics experiments and has been recently revived in the context of string theories at the Grand Unified Theory scale. An exponentially growing density of states causes the final explosion of the black hole to be much more violent. Once a limiting temperature is reached, the remaining mass is suddenly emitted into all available states. This occurs when the exponential increase in the number of states offsets the exponential

decrease in the tail of the Hawking radiation at a critical black-hole mass of $M_c = (1.06 \times 10^{13} \text{ GeV})/\Lambda$ g. More realistically, equations of the form of equation (6) may indicate that a phase transition, whose details are as yet unknown, occurs at Λ . In this case, the quasi-exponential spectrum would not continue until $m \rightarrow \infty$ but would last for, at most, the duration of the finite phase transition. Again we could observe bursts carrying off a significant fraction of the PBH's mass from holes just reaching the phase transition temperature. In the final section, we shall calculate the chance of detecting the final evaporation stage both for a Hagedorn scheme with $\Lambda = 160$ MeV and for the standard model. These cases represent two extreme extrapolations of the particle model.

Diffuse backgrounds from Hawking radiation

By integrating equation (4), one can relate²¹ the initial mass M_i of the black hole to its lifetime, that is, the time, τ_{evap} , over which it completely evaporates. To a good approximation we have

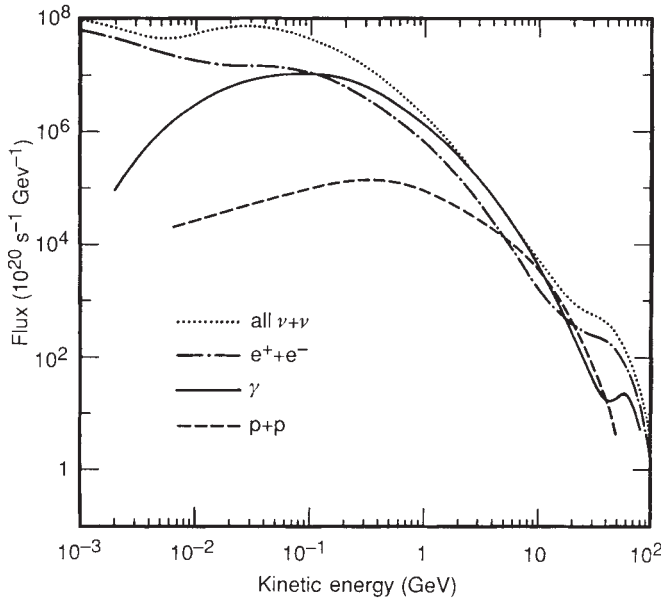
$$M_i \approx [3\alpha(M_i)\tau_{\text{evap}}]^{1/3} \\ \approx 5.4 \times 10^8 \left[\frac{\alpha(M_i)}{5.3 \times 10^{25} \text{ g}^3 \text{ s}^{-1}} \right]^{1/3} \left[\frac{\tau_{\text{evap}}}{1 \text{ s}} \right]^{1/3} \text{ g} \quad (7)$$

A hole that formed in the early Universe and is just completing its evaporation today therefore had an initial mass $M_* \approx [3\alpha(M_*)t_0]^{1/3}$ where t_0 is the age of the Universe. In Friedman models, t_0 depends on the relative matter density Ω_m and the Hubble constant $H_0 = h_0/100 \text{ km s}^{-1} \text{ Mpc}^{-1}$. Because M_* is close to the thresholds for $\mu^\pm$, $\pi^{0,\pm}$ and u, d quark emission, we integrate equation (4) and solve numerically for M_* by approximating α linearly at each threshold²⁹. We find that M_* lies in the range¹⁴

$$4.3 \times 10^{14} < M_* < 7.0 \times 10^{14} \text{ g} \quad (8)$$

for $125 \leq \Lambda_{\text{qh}} \leq 300$ MeV, $0.06 \leq \Omega_m \leq 1.0$ and $0.4 < h_0 < 1$. As $\tau_{\text{evap}} \propto M_i^3$, black holes with $M_i \gg M_*$ have not lost a significant amount of mass by today, whereas those with $M_i < M_*$ have completely evaporated.

PBHs could have formed in the early Universe in a number of ways¹. They may, for example, have been produced by initial density inhomogeneities² or as a result of a phase transition³. Hawking, Zembowicz and Polnarev have also recently proposed that a spectrum of black holes may have been created by the collapse of cosmic strings^{4,5}. If spherically symmetric PBHs form from scale-invariant initial density perturbations that have a


 FIG. 2 Particle fluxes radiated by a black hole of temperature $T=10$ GeV.

gaussian distribution, the number density of holes created with mass in the range $(M_i, M_i + dM_i)$ is³⁰:

$$\frac{dn}{dM_i} = \frac{\mathcal{N}}{M_*} \left[\frac{M_i}{M_*} \right]^{-\beta} \quad (9)$$

Here $\mathcal{N} = (\beta - 2)\Omega_{\text{PBH}}\rho_c / M_*$, where Ω_{PBH} is the present fraction of the cosmological critical mass density ρ_c in holes of $M \geq M_*$. $\mathcal{N}$ corresponds to the initial number density of holes per logarithmic mass interval at $M = M_*$. If the equation of state of the Universe is $p = \gamma p$ at the formation epoch, where p is pressure, then $\beta = (1 + 3\gamma)/(1 + \gamma) + 1$. For PBHs forming in the radiation-dominated era, we have $\beta = 2.5$.

The emission from PBHs over the lifetime of the Universe will produce diffuse particle backgrounds. These can be compared with observations to obtain an upper limit¹¹ on Ω_{PBH} and hence $\mathcal{N}$. The predicted diffuse photon spectrum is shown in Fig. 3 for the initial distribution equation (9) with $\beta = 2.5$. We have included quark and gluon emission and neglected all losses other than redshift. The integrated emission²⁹ shown in Fig. 3 has an E^{-1} slope for photon energies well below 100–300 MeV, coming from jet fragmentation, and turns over into an E^{-3} slope above 300 MeV. The E^{-3} slope, which is independent of the value of β , is dominated by direct photon emission and comes from evaporation in the present epoch. Matching to the observed γ -ray spectrum at 100 MeV (where photon losses are negligible), as was done in Fig. 3, yields a bound on the density of PBHs¹²

$$\Omega_{\text{pbh}} \leq 7.6(\pm 2.6) \times 10^{-9} h_0^{-1.95 \pm 0.15} \quad (10)$$

for $\Omega_m = 1.0$ and $\beta = 2.5$. This corresponds to $\mathcal{N} \leq 10^4 \text{ pc}^{-3}$. The limit is $\sim 60\%$ weaker for $\Omega_m = 0.06$ and varies by $\lesssim 50\%$ when the value of Λ_{qh} is changed. MacGibbon and Carr¹² also obtain bounds on Ω_{pbh} from the antiproton, electron and positron fluxes. These bounds are remarkably similar to the γ -ray limit (equation (10)) if the PBHs are clustered to the same degree as other material in the galactic halo. (In this case, the local density is enhanced by a large factor which we estimate below to be in excess of 10^7 .) The bounds may even overlap, allowing the possibility that PBH emission may be contributing significantly to all observed interstellar cosmic ray and γ -ray spectra between 0.1 and 1 GeV. It should be pointed out that none of these conclusions depend sensitively on the initial distribution (equation (9)). The limit given in equation (10) is always valid, as it is essentially determined by the peak emission from the holes with critical mass M_* (see equation (3)).

From the initial PBH distribution, we can calculate the distribution of holes at any later time t . If we invoke equation (7), the distribution of equation (9) becomes¹⁰

$$\begin{aligned} \frac{dn}{dM} &= \frac{dn}{dM_i} \frac{dM_i}{dM} \\ &\approx \frac{\mathcal{N}}{M_*} \left[1 + \left(\frac{t}{t_0} \right) \left(\frac{\alpha(M)}{\alpha(M_*)} \right) \left(\frac{M}{M_*} \right)^{-3} \right]^{-(\beta+2)/3} \\ &\quad \times \left[\frac{M}{M_*} \right]^{-\beta} \left[\frac{\alpha(M)}{\alpha(M_i)} \right]^{(\beta-1)/3} \\ &\propto \begin{cases} M^2, & M \ll 10^9 t^{1/3} \text{ g} \\ M^{-\beta}, & M \gg 10^9 t^{1/3} \text{ g} \end{cases} \end{aligned} \quad (11)$$

where M is the mass at time t of a hole with initial mass M_i . Note that the present distribution for $M \ll M_*$ is independent of β . Regardless of their formation mechanism, holes today with masses less than M_* have the distribution $dn/dM \propto M^2$. From equation (11), or directly from equation (4), we now obtain the current rate of expiring PBHs,

$$\frac{dn}{dt} = \frac{\alpha(M_*)}{M_*^3} \mathcal{N} \approx \frac{\mathcal{N}}{t_0} \quad (12)$$

Allowing for uncertainties in M_* and $\alpha(M_*)$, the present rate of explosions averaged over the Universe should thus lie in the range

$$\begin{aligned} 1.2 \times 10^{-11} \text{ yr}^{-1} (h_0 = 0.4) \\ < \frac{dn}{dt} \mathcal{N}^{-1} \\ < 4.0 \times 10^{-11} \text{ yr}^{-1} (h_0 = 1.0) \end{aligned} \quad (13)$$

As they should not have large peculiar velocities, one should expect PBHs to cluster in galactic haloes along with other material. This will enhance their local density by a factor ζ . Assuming an isothermal halo we estimate² a local enhancement factor at $R_\odot$, the Sun's distance from the Galactic Centre, of

$$\zeta = 1.36(\pm 0.9) \times 10^7 \left(\frac{\Omega_\odot}{0.01} \right)^{-1} h_0^{-2} \quad (14)$$

where $\Omega_\odot$ is the visible fraction of ρ_c in galaxies out to $R_\odot$. A detailed discussion will be given elsewhere (F.H., J.H.M., T.C.W. and E.Z., manuscript in preparation). The large uncertainties associated with clustering will strongly affect our chances for detecting PBH evaporations. For $\mathcal{N} = 10^4 \text{ pc}^{-3}$, the present local rate of explosions could be as low as $10^{-7} \text{ pc}^{-3} \text{ yr}^{-1}$, if holes are not clustered, or as high as $10 \text{ pc}^{-3} \text{ yr}^{-1}$ if holes are clustered. As γ -rays (unlike charged particles) are not confined in the Galaxy, the diffuse photon background from PBHs is not affected by clustering.

Final-stage emission

The final radiation from a black hole is dictated by particle physics. We will first assume the degrees of freedom of the particle to be those of the standard model, even at very high temperatures. Let us consider a black hole of mass M_B which completely evaporates in a time $\Delta\tau$ given by equation (7). Photons are emitted either directly by the Hawking mechanism of equation (3) or as fragmentation products of quarks or gluons (partons). The quark flux, integrated over $\Delta\tau$, peaks at an energy of

$$Q \approx 40 \left[\frac{1 \text{ s}}{\Delta\tau} \right]^{1/3} \text{ TeV} \quad (15)$$

Calculation of the photon flux resulting from fragmentation of these quarks is straightforward and proceeds in three stages:

one starts with the parton fluxes, which fragment into neutral pions, and these then decay into two photons. Here we will follow ref. 31 and choose an empirical fragmentation function of the form $dN_\pi/dz = \frac{1}{16}z^{-3/2}(1-z)^2$, where $z = E_\pi/E_{\text{jet}}$.

The predictions relevant to experiment are obtained by integrating the photon flux above a detector threshold E_D . If the lifetime of the hole is short compared with the observation time, one must integrate its emission spectrum over its history. Assuming that the flux before decay is emitted near the peak energy Q in the Hawking distribution, we find that the time-integrated photon flux from both direct and fragmentation photons decreases as E_γ^{-3} above $E_\gamma \approx Q$. This is a consequence of the time dependence of the hole's temperature given by equations (2) and (7). Below $E_\gamma \approx Q$, the flux traces the slope of the fragmentation function, peaks near the pion mass and is eventually cut off at $E < m_\pi$. Decay photons dominate at all energies, because there are 72 quark and 16 gluon degrees of freedom, but only two direct photon degrees. The contribution per helicity state is also greater for $s = \frac{1}{2}$ than for $s = 1$ states. The number of photons emitted above E_D is

$$N_\gamma(>E_D) \approx 2.4 \times 10^{37} \left(\frac{\text{GeV}}{Q} \right)^2 \times \begin{cases} \left[\sqrt{\frac{Q}{E_D}} \left(\frac{5}{6} + \frac{3E_D}{Q} + \frac{5E_D^2}{14Q^2} \right) - \frac{5E_D}{3Q} - \frac{5}{2} \right] + \frac{1}{250}, & E_D < Q \\ \left(\frac{Q}{E_D} \right)^2 \left[\frac{1}{42} + \frac{1}{150} \left(1 - \frac{2}{5} \max \left(0, 2 - \frac{Q}{E_D} \right) \right) \right], & E_D \geq Q \end{cases} \quad (16)$$

The direct photons contribute the last terms in these expressions.

Alternatively, if the lifetime is large compared with the observation time, we can freeze the spectrum and calculate the instantaneous flux neglecting the evolution of the hole

$$\frac{dN_\gamma}{dt}(>E_D) \approx 8.0 \times 10^{23} \frac{Q}{1 \text{ GeV}} \times \begin{cases} \left[\frac{1}{4} \left(\frac{Q}{E_D} \right)^{1/2} \left(\frac{1+E_D^2}{Q^2} \right) - 1 \right] + \frac{3}{2} \left(\frac{E_D}{Q} \right)^{1/2} - \frac{E_D}{Q} \Big] s^{-1} & (17) \end{cases}$$

The detailed structure of the above expressions reflects our choice of fragmentation function.

At energies close to the black-hole temperature, other processes such as prompt photon production may also enhance the flux. The neutrino and cosmic-ray emission from the PBHs is comparable to the photon emission¹³. But as the neutrino and cosmic-ray backgrounds are greater than any postulated diffuse photon background, photon detectors offer the greatest sensitivity to search for individual evaporating PBHs.

Observing isolated black holes

Can the Hawking radiation of an individual hole be detected above the backgrounds created by the distribution of evaporating PBHs or above the observed extragalactic background? If n_{bh} is the local number density of PBHs with present mass less than M_{th} , then integration of equation (9) yields

$$n_{\text{bh}} \approx \zeta \mathcal{N} \frac{\alpha(M_*)}{3\alpha(M_{\text{th}})} \left[\frac{M_{\text{th}}}{M_*} \right]^3 \quad (18)$$

for $M_{\text{th}} \ll M_*$. According to the mass distribution of equation (18), we expect the closest hole in a volume n_{bh}^{-1} to have an average mass of $0.75 M_{\text{th}}$. The photon flux at the detector from

this hole is

$$\frac{d^3 N_\gamma}{dA dt dE_\gamma} = \frac{d^2 N_\gamma}{dt dE_\gamma} \left[\frac{M_{\text{th}}}{M_*} \right]^2 \left[\frac{\alpha(M_*)}{\alpha(M_{\text{th}})} \right]^{2/3} \frac{\mathcal{N}^{2/3}}{3\sqrt[3]{12\pi}} \quad (19)$$

where $d^2 N_\gamma/dt dE_\gamma$ is the instantaneous flux computed for direct photons according to equation (3) and for parton fragmentation photons according to equation (17).

Let us turn first to the diffuse background created by the total emission of all evaporating black holes. Using the results quoted in the previous section, we can represent the PBH background by²⁹

$$\frac{d^4 N_\gamma}{dA dt d\omega dE_\gamma} \approx 7 - 14 \times 10^{-11} \left[\frac{\mathcal{N}}{1 \text{ pc}^{-3}} \right] E^{-3} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ GeV}^{-1} \quad (20)$$

above 300 MeV, where ω is solid angle, for $0.06 \leq \Omega_m \leq 1.0$, $0.3 \leq h_0 \leq 0.7$ and $\beta = 2.5$. Second, the extragalactic diffuse photon background measured by satellites^{33,34,35} between 35 and 170 MeV (ref. 14) is

$$\frac{d^4 N_\gamma}{dA dt d\omega dE_\gamma} = 2.7(\pm 0.5) \times 10^{-4} \left[\frac{E_\gamma}{100 \text{ MeV}} \right]^{-2.4(\pm 0.2)} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \text{ GeV}^{-1} \quad (21)$$

Observations above 170 MeV have not yet been made. The EGRET instrument, recently launched, will be capable of detecting photons up to energies of ~ 20 GeV. As the galactic emission falls off less steeply than the extragalactic background, however, it is not known if the extragalactic diffuse background can be resolved from the galactic emission above a few hundred MeV, even at high galactic latitude. Regardless of their source, photons with energies $> 10^6$ GeV are cut off by pair production on the cosmic microwave background.

To estimate the detector resolution necessary to resolve a point source in the presence of a diffuse background, we require

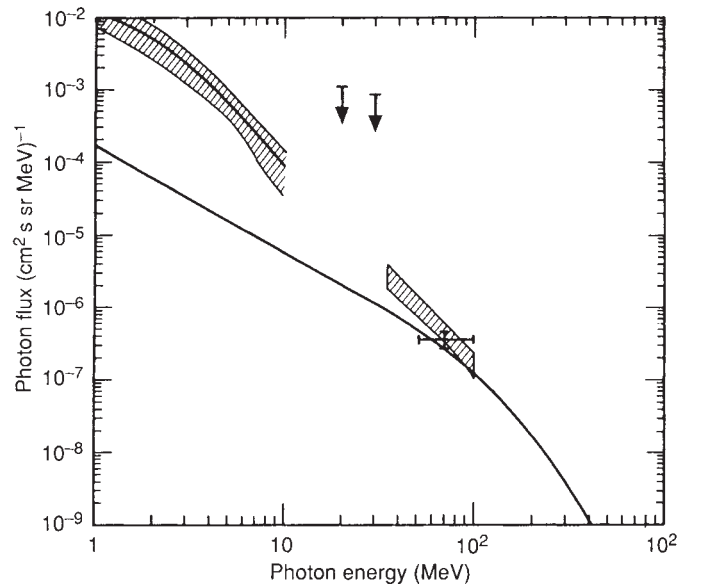


FIG. 3 Diffuse photon flux from Hawking radiation by a distribution of black holes. Shown as a solid line is the calculation for the black-hole density $\Omega_{\text{pbh}} = 7.6 \times 10^{-9}$ leading to the bound in equation (10). Any concentration in excess of this limit would exceed the observed diffuse γ -ray background. Data and upper limits from ref. 28. The hatched areas correspond to 1σ errors.

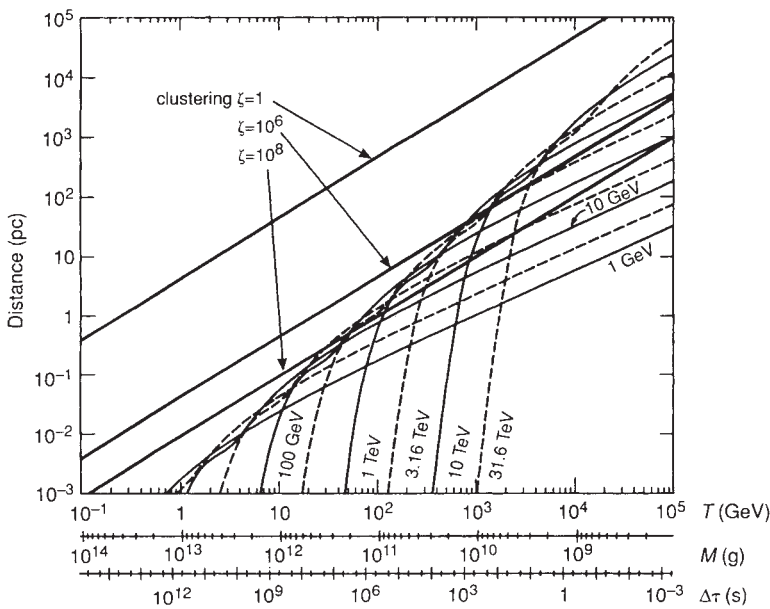


FIG. 4 The three straight lines show the distance at which the closest black hole of temperature $> T$ is expected assuming the maximum density allowed by the bound in Fig. 3. The result is shown for three values of the clustering factor ζ . The other lines show the distance at which the flux from such black hole can be resolved from the diffuse background. Detectors are assumed to have an angular resolution $\omega_D = 10^{-3}$ sr and the energy thresholds shown. The MeV bound in Fig. 3 can be improved by a detector of a given threshold in the region where its sensitivity curve exceeds the straight line.

that the diffuse photon flux within the resolved solid angle of the detector ω_D does not exceed the point source flux; that is,

$$\frac{d^3 N_\gamma}{dA dt dE_\gamma} \geq \omega_D \frac{d^4 N_\gamma}{dA dt d\omega dE_\gamma} \quad (22)$$

As a first example, let us consider the direct photon emission and substitute equations (19) and (20) into (22), assuming $\omega_D = 10^{-3}$ and $\zeta N = 10^{10}$. In this case we find that black holes of mass $M < 4 \times 10^{10}$ g and peak photon energy $E_\gamma \approx (8.66 \times 10^{13} \text{ g})/M_{\text{th}}^{-1} \text{ GeV}$ will stand out from the PBH background above ~ 2 TeV. If we compare equation (20) instead with the extrapolation of the measured background, equation (21), the critical energy turns out to be about two orders of magnitude greater. Alternatively, if we look at photon fragments of jets with $E_\gamma/E_{\text{jet}} \approx 0.1$, for example, we find that both critical energies are about two orders of magnitude smaller. Thus it may be possible to resolve the instantaneous decay emission from the background at energies above 0.01–1 TeV. This conclusion depends strongly on the nature of the fragmentation function as $E_\gamma/E_{\text{jet}} \rightarrow 1$ and on the degree to which the PBHs cluster in the galactic halo.

For an estimate more relevant to detectors, we integrate the individual PBH flux above the detector threshold E_D (see equation (16)). Our conclusions are summarized in Fig. 4, which shows the distances at which a hole of temperature T produces a flux in the detector matching the maximum diffuse PBH background allowed by the 100-MeV data. The direct and fragmentation photons produce the bumps at high and low energy, respectively. If the PBHs are clustered with $\zeta N = 10^{12}$, then detectors with $E_D \geq 30$ GeV may be able to resolve holes of $T \geq 20$ GeV from the background. If $\zeta N = 10^{10}$, however, the energy threshold for the detector must exceed 300 GeV and the black-hole temperature must exceed 700 GeV. For low temperatures, the detector should see a sky uniformly bright in black holes, and we have a version of Olber's paradox for black holes. In this case, the black-hole horizon is extremely small and the radiation density of the background sky is much smaller than the black-hole surface radiation density. No practical detector can single out a point source because a large number of sources appear in any field of view.

Although the results are model-dependent, the qualitative features illustrated in Fig. 4 should be correct. The distance to the nearest hole of a given temperature should scale as $\zeta^{-1/3}$. The curves shift upward as $\sqrt{10^{-3}/\omega_D}$ with improved detector resolution. Optimum signal-to-noise ratio is obtained at $E_D \approx T$. In Fig. 4, we have used an E^{-3} diffuse background produced

by a PBH distribution. If the observed background falls as $E^{-2.4}$, as extrapolated from the MeV data, the conclusions are more restrictive.

At the higher energies, air-shower arrays can search for a diffuse photon background, for example by exploiting the small muon abundance in photon-induced air showers to reject cosmic rays³². From equation (20) we expect a flux of

$$\frac{d^2 N_\gamma}{dA dt} (> 100 \text{ TeV}) < 3 \times 10^{-16} \text{ cm}^{-2} \text{ s}^{-1} \quad (23)$$

from the PBH background. This flux is unfortunately below the flux expected from galactic cosmic rays interacting with the interstellar medium (ref. 19; and F.H., J.H.M., T.C.W. and E.Z., in preparation). Note, however, that the lifetime of a 10 TeV black hole is ~ 1 s (see equation (7)), and therefore our previous discussion of the final stages of black holes is also relevant here.

Detection and bounds

Ground-based γ -ray telescopes, which detect air showers initiated by photons of energy TeV to PeV (see Box 2), can be used to search for PBHs or to set limits on their density expiring in the present epoch. A revitalized search for PBHs has been prompted by the development of new atmospheric Čerenkov and air-shower array telescopes which can identify γ -ray showers amid the much greater background of charged cosmic rays, and by the more definite particle physics predictions for final evaporation discussed in the previous section. The proposed detection techniques are generally based on the same principles as before. Experimentally there are two fundamental considerations in such searches. First, how large a volume can the telescope search efficiently and for what exposure time? Second, what conclusively distinguishes individual PBH emissions from a chance coincidence of background cosmic-ray showers?

At 100 TeV energies, experiments are being commissioned with collection areas in excess of 10^{10} cm^2 and angular resolutions less than 4×10^{-4} sr (ref. 35). Hadron rejection ratios,

TABLE 1 Summary of γ -ray flux at energies above E_D

$\Delta\tau$	10^6 s	10^3 s	1 s
$N_\gamma(>E_D, \tau < \Delta\tau)$	4.4×10^{31}	4.8×10^{30}	2.2×10^{29}
$\frac{dN_\gamma(>E_D)}{dt}$	$1.0 \times 10^{25} \text{ s}^{-1}$	$2.5 \times 10^{27} \text{ s}^{-1}$	$1.3 \times 10^{29} \text{ s}^{-1}$

TABLE 2 Observational limits on PBH densities

Technique*	Group†	Assumptions‡	Density of events (pc ⁻³ yr ⁻¹)	References	
ACT	3 × 1.5 m	SAO-UCD	HM	<0.47	Ref. 44
ACT	10 m	SAO-UCD	HM	<0.67	Ref. 44
ACT	4 × 1.5 m	SAO-UCD	HM	<2.1	Ref. 45
ACT	10 m + 9 m	SAO-UCD	HM	<0.04	Ref. 45
ACT	10 m	SAO-UCD	EPM	<7 × 10 ⁵	Ref. 45
ACT	10 m	SAO-UCD	EPM	<8 × 10 ⁵	Ref. 46
ACT	2 × 1.5 m	SAO-UCD	EPM	<3 × 10 ⁴	Ref. 46
ACT	10 m	SAO-UCD	EPM	<2 × 10 ⁴	Ref. 47
EAS	2 arrays	UCC-UCD	EPM + UHE emission	<2 × 10 ⁴	Ref. 48
EAS	1 array	Tata	EPM + UHE emission	<3 × 10 ³	Ref. 49
Optical	3 × 1.5 m	UCD	EPM + optical emission	<0.3	Ref. 50
Radio	10 m	SAO	EPM + radio emission	<3 × 10 ⁻⁷	Ref. 51
Radio			EPM + radio emission	<2 × 10 ⁻⁷	Ref. 52
Radio	Areicho	U. Mass	EPM + radio emission	<7 × 10 ⁻¹⁰	Refs 53-55

* ACT: atmospheric Čerenkov technique; EAS: extensive air shower.

† SAO: Smithsonian Astronomical Observatory; UCD: University College Dublin.

‡ HM: Hagedorn model; EPM: elementary particle model or standard model; UHE: ultra-high-energy.

based on the muon-to-electron ratio, of 10^{-4} – 10^{-5} (ref. 37) are expected, although this has yet to be demonstrated in the detection of a discrete source. Air-shower arrays are now being discussed with lower energy thresholds in the 10–100 TeV range^{37,38}. As an illustration, let us consider a somewhat idealized air-shower array telescope with the following parameters: collection area $A_D = 10^8 \text{ cm}^2$, angular resolution $\omega_D = 4 \times 10^{-4} \text{ sr}$, energy threshold $E_D = 10 \text{ TeV}$, operation time 1 year and hadron rejection ratio 10^{-3} . We will assume a typical zenith angle cutoff of 30° , and hence an acceptance angle of $\omega_A = 0.75 \text{ sr}$. The depth of sky covered by the telescope can be calculated from the minimum flux sensitivity for detection of a burst. If n_γ is the minimum number of γ -rays of energy $> E_D$ required to register as a burst, then

$$n_\gamma = \frac{A}{4\pi r^2} N_\gamma(>E_D) \quad (24)$$

where r is the distance to the PBH. From equation (7), both extensive air-shower arrays and Čerenkov telescopes become sensitive to the direct photon flux from a black hole at a time

$$\Delta\tau \leq \left[\frac{5 \times 10^{13} \text{ g GeV}}{E_D} \right]^3 \frac{1}{3\alpha(M)} \quad (25)$$

before the hole completely evaporates. The relevant emission period, $\Delta\tau$, is very sensitive to the detector threshold E_D . If $E_D = 10 \text{ TeV}$, then $\Delta\tau \approx 64 \text{ s}$, which in turn fixes the black-hole temperature and the peak photon energy Q for this detector. The hole emits a total flux of $\sim 5.7 \times 10^{27}$ photons above threshold in the last 74 s; see equation (16). Of this flux, 77% is emitted in the last second. If we require a signal of n_γ photons within $\Delta\tau = 1 \text{ s}$, and within the detector resolution, the depth of sky scanned by the detector is

$$r_\gamma \approx 0.035 \sqrt{\frac{A}{10^8 \text{ cm}^2} \frac{3}{n_\gamma}} \text{ pc} \quad (26)$$

For the above values, $r_\gamma = 0.035 \text{ pc}$, giving a scanned volume of $V_\gamma = \omega_A r^3/3 = 1.1 \times 10^{-5} \text{ pc}^3$.

The relevant background for our search comes from hadronic cosmic rays with energy close to E_D . The background rate of cosmic-ray showers is

$$\frac{d^3 N_{\text{cr}}}{dA dt d\omega}(>E_D) = 0.21 \left[\frac{E}{1 \text{ GeV}} \right]^{-1.5} \text{ cm}^{-2} \text{ s}^{-1} \text{ sr}^{-1} \quad (27)$$

For $E_D = 10 \text{ TeV}$ and $\omega_D = 10^{-3} \text{ sr}$, the telescope would see a background of 0.02 s^{-1} per resolution element. If the rejection

ratio is 10^{-3} , the minimum detectable photon event rate is then $2 \times 10^{-5} \text{ s}^{-1}$ per resolution element. For $n_\gamma = 3$ and $\Delta\tau = 1 \text{ s}$, the chance of cosmic rays mimicking individual PBH emission from any direction is $< 10^{-4} \text{ yr}^{-1}$. If no bursts were detected in one year of operation, we could set a limit of 4.3 events yr^{-1} at the 99% confidence level in V_γ . This would lead to an upper limit on the PBH explosion density of $dn/dt \leq 3.9 \times 10^5 \text{ yr}^{-1} \text{ pc}^{-3}$.

An alternative approach is to use atmospheric Čerenkov telescopes. These have considerably lower E_D and much smaller fields of view than air-shower arrays and must be operated under clear dark skies. Sophisticated detectors can reject background hadronic cosmic rays with 99% efficiency using differences in the Čerenkov images from hadronic and photonic air showers³⁹. To evaluate the PBH detection efficiency, we consider the GRANITE telescope⁴⁰, now under construction at the Whipple Observatory. It consists of two 10-m-aperture optical concentrators, each fitted with a 109-pixel electronic camera, and separated by 120 m. The telescope is triggered by coincidences between corresponding pixels in the two concentrators. The energy threshold for γ -ray shower detection is 100 GeV, the full field is 2.5° ($\omega_A = 1.4 \times 10^{-3} \text{ sr}$), the angular resolution is of the order of the pixel size, 0.25° , and the collection area is $A_D = 5 \times 10^8 \text{ cm}^2$.

Lower thresholds imply sensitivity to the larger fluxes resulting from jet fragmentation, and hence larger probed depths. This compensates for the reduced field of view so that their sensitivity is comparable to extensive air-shower arrays. Table 1 summarizes the flux and time-integrated flux, above E_D , as estimated from equations (16) and (17). For example, for small zenith angles and $\Delta\tau = 1 \text{ s}$, the probed depth is $r_\gamma \approx 0.48 \text{ pc}$ and the sensitive volume is $V_\gamma \approx 5.2 \times 10^{-5} \text{ pc}^3$. The sensitivity can be further improved by exploiting the unusual zenith angle variation of Čerenkov telescopes. Simulations⁴¹ show that as zenith angle increases, E_D and A_D increase by a factor f . At a zenith angle of 75° , we have $f = 100$. In this case, $A_D = 5 \times 10^{10} \text{ cm}^2$ and $E_D = 10 \text{ TeV}$, which, for $n_\gamma = 4$, gives $r_\gamma \approx 0.72 \text{ pc}$ and $V_\gamma \approx 1.7 \times 10^{-4} \text{ pc}^3$.

The background of cosmic-ray events comes from distant proton showers of similar energies which arrive nearly parallel to the telescope's optical axis and from more local showers which traverse both telescopes' fields of view and radiate into the cone of the cameras. The former is estimated at $0.0015 \text{ s}^{-1} \text{ pixel}^{-1}$ for a rejection ratio of 10^{-2} . The latter can be rejected by examining the images. An observing time of 1,000 h in a dedicated experiment and $n_\gamma = 3$ will give a chance of < 0.01 of mimicking a PBH. If no events are found in this time, then an upper limit on the number of PBH explosions of $\leq 1.3 \times 10^5 \text{ yr}^{-1} \text{ pc}^3$ for $V_\gamma = 3 \times 10^{-4} \text{ pc}^3$ can be inferred. This is weaker

than the limit obtained from the MeV bound assuming maximum clustering given in equation (14).

In the alternative model with an exponentially growing number of degrees of freedom, 6.0×10^{34} erg of energy is evaporated in a final burst lasting 10^{-7} s. A significant fraction of the emission will be γ -rays of average energy 250 MeV. In this case the volume scanned by a dedicated telescope is large enough to strengthen the MeV bound on Ω_{PBH} considerably. The most sensitive detectors are atmospheric Čerenkov telescopes which detect bursts as if they were large air showers. A burst of 250-MeV γ -rays strikes the top of the atmosphere like a plane wave and interacts to produce secondary electrons. These in turn cause the atmosphere to radiate Čerenkov light which reaches the detector. Several independent estimates suggest that each vertically incident 250-MeV γ -ray produces on average $\sim 2,300$ optical photons reaching detector level^{37,42}. Although individual photons cannot be detected, the combined effect is equivalent to a giant shower of 100 ns duration. The challenge is to design a detector that can detect the weakest possible γ -ray flux (and hence be sensitive to the greatest possible depth) while at the same time maximizing the field of view (and hence the volume V_γ). The accidental trigger rate from background events (such as local air showers, night-sky fluctuations and man-made sources) must be minimized to get maximum exposure time under clear, dark, night-sky conditions. In general, these conditions are not found in conventional atmospheric Čerenkov experiments, and the most sensitive measurements require specially designed experiments.

The most sensitive direct limit on the explosion rate (Box 3) was set by the operation of two large optical reflectors in coincidence in 1976 (ref. 42). In this experiment a cluster of phototubes was operated in coincidence with a 9-m optical reflector at the White Sands Missile Range and on a 10-m reflector at the Whipple Observatory, 400 km away. Both reflectors were pointed in the same direction under clear skies for 22.5 h; no coincidences were seen. The threshold for the detection of an event was 26 optical photons m^{-2} , giving a burst sensitivity of 5.6×10^{-10} erg cm^{-2} . For $E_B \approx 10^{34}$ erg, this gave a maximum detectable distance of 390 pc, a sensitive volume of $V_\gamma =$

$4.4 \times 10^4 \text{ pc}^3$ and in this model a limit on bursts of $< 0.04 \text{ pc}^{-3} \text{ yr}^{-1}$.

Greater sensitivity could be achieved using the two 11-m aperture Solar concentrators at Sandia Labs, Albuquerque⁴³ together with the GRANITE telescopes ($2 \times 10 \text{ m}$)⁴⁰ now under construction at the Whipple Observatory in Arizona. Each reflector would be equipped with an array of phototubes to give a full field of view of 3° . A local coincidence would be demanded between each set of telescopes which would then be put in coincidence (by telephone link) with each other. The 100-ns duration of the pulses would be verified using waveform digitizers at each telescope (eliminating possible spurious events). With the increased collection area and a more sophisticated trigger, the threshold per burst could be reduced to 1.8×10^{-10} erg cm^{-2} , increasing V_γ by a factor of eight. The exposure time could also be increased to 450 h in six months of operation. A PBH limit as low as 2.5×10^{-4} explosions $\text{pc}^{-3} \text{ yr}^{-1}$ might then be obtained.

Results

Within the context of the standard model of quarks and leptons, we have discussed the signatures of black-hole evaporations and obtained a firm bound on the PBH abundance from MeV data. The standard model is incomplete, and new degrees of freedom should be revealed at yet higher energies. We have argued the possibility for improving the MeV bound by exploiting the new generation of TeV and PeV telescopes. Even in the absence of a detection, improvements of the bounds are a worthwhile endeavour as they provide direct information about the homogeneity and isotropy of the early Universe. $\square$

F. Halzen and E. Zas are at the Department of Physics, University of Wisconsin, Madison, Wisconsin, USA. J. H. MacGibbon is at Code 665 NASA/Goddard Space Flight Center, Greenbelt, Maryland 20771, USA. T. C. Weekes is at the Whipple Observatory, Harvard Smithsonian Center for Astrophysics, P.O. Box 97, Amado, Arizona 85645-0097, USA.

BOX 3 Previous searches for PBH γ -ray bursts

Several direct experimental searches for PBH explosions were made in the 1970s following the theoretical discovery of PBH evaporation. In general, these involved the use of ground-based γ -ray telescopes and used both archival data and observations made with specially configured detectors. No statistically significant detection was reported, but upper limits were derived that placed limits on the PBH density.

Two extreme models for the final stages of the evaporation were considered, each predicting quite different parameters for the γ -ray burst. The experiments were confined to searches for bursts with these characteristics. The standard model (also known as the elementary particle model or EPM) predicted an outburst of 10^{30} erg in γ -rays of energy 5 TeV, lasting 0.1 s. The Hagedorn model, which corresponds to the exponential model discussed in the text, predicted an outburst of 10^{34} erg in 250 MeV γ -rays, lasting 100 ns.

In addition, it was postulated⁵⁰⁻⁵⁵ that under certain conditions the exploding PBH might produce detectable radio or optical emission. In this scenario, the final e^+e^- emission from the PBH is braked by an ambient magnetic field as the particle shell expands relativistically away from the hole, thus producing an electromagnetic pulse. The pulse would only be appreciable if the explosion is governed by the Hagedorn model—a conclusion strengthened by the most recent analysis including quark and gluon emission¹². Because of the greater sensitivity of radio and optical techniques, it was possible to probe greater depths with these experiments. A null result in these searches had less force, however, as the conditions for pulse production are highly model-dependent.

In Table 2 we list the principal searches that have been made, the techniques and assumptions used and the limits derived on the number density of PBHs; see refs 44-55. The latter show a spread of 10^{14} .

1. Carr, B. J. in *Observational and Theoretical Aspects of Relativistic Astrophysics and Cosmology* (eds Sanz, J. L. & Goicoechea, L. J.) 1-78 (World Scientific, Singapore, 1985).
2. Hawking, S. W. *Mon. Not. R. astr. Soc.* **152**, 75-78 (1971).
3. Hawking, S. W. *et al. Phys. Rev. D* **26**, 2681-2693 (1982).
4. Hawking, S. W. *Phys. Lett.* **B231**, 237-239 (1989).
5. Polnarev, A. & Zembowicz, R. *CAMK preprint 194* (1988).
6. Hawking, S. W. *Nature* **248**, 30-31 (1974).
7. Novikov, I. D. *et al. Astr. Astrophys.* **80**, 104 (1979).
8. Zeldovich, Ya. B. & Starobinskiĭ, A. A. *Sov. Phys. JETP* **24**, 571 (1976).
9. Hall, S. J. & Hsu, S. *Berkeley preprint LBL-27982* (1989).
10. Carr, B. J. *Astr. J.* **206**, 8-25 (1976).
11. Page, D. N. & Hawking, S. W. *Astrophys. J.* **206**, 1-7 (1976).
12. MacGibbon, J. H. & Carr, B. J. *Astrophys. J.* **371**, 447-469 (1991).
13. MacGibbon, J. H. & Webber, B. R. *Phys. Rev. D* **41**, 3052-3079 (1990).
14. MacGibbon, J. H. *Phys. Rev. (in the press)*.
15. Weekes, T. C. *Phys. Rep.* **160**, 1-121 (1988).
16. Bonnet-Bidaud, J. M. & Chardin, G. *Phys. Rep.* **170**, 325 (1988).
17. Protheroe, R. in *Proc. 20th ICRC, Moscow, 1987*, vol. 8 (ed. Kozyrivsky, V. A. *et al.*) (Nauka, Moscow, 1987).
18. Nagel, D. E., Gaisser, T. K. & Protheroe, R. J. *Ann. Rev. Nucl. Part. Sci.* **38**, 609-657 (1988).
19. Halzen, F., Protheroe, R. J., Stanev, T. & Vankov, H. P. *Phys. Rev. D* **41**, 342-346 (1990).
20. Halzen, F., Protheroe, R. J., Stanev, T. & Vankov, H. P. in *Proc. 21st ICRC, Adelaide* (ed. Protheroe, R. J.) (Adelaide University, 1990).
21. Page, D. N. *Phys. Rev. D* **13**, 198-206 (1976).
22. Page, D. N. *Phys. Rev. D* **16**, 2402-2411 (1977).
23. Elster, T. *J. Phys. A* **16**, 989-996 (1983).
24. Elster, T. *Phys. Lett. A* **92**, 205-209 (1983).
25. Spinkins, R. D. thesis, Pennsylvania State Univ. (1986).
26. Oliensis, J. & Hill, C. T. *Phys. Lett. B* **143**, 92-102 (1984).
27. Huang, K. & Weinberg, S. *Phys. Rev. Lett.* **25**, 895-897 (1970).
28. Hagedorn, R. *Nuovo Cim.* **LVIX**, No. 4, 1027-1057 (1968).
29. MacGibbon, J. H. thesis, Univ. of Cambridge (1988).
30. Carr, B. J. *Astrophys. J.* **201**, 1-19 (1975).
31. Hill, C. T., Schramm, D. V. & Walker, T. P. *Phys. Rev. D* **36**, 1007-1016 (1987).
32. Halzen, F., Stanev, T., Zas, E. & Gaisser, T. K. in *Proc. 21st ICRC, Adelaide*, vol. 9 (ed. Protheroe, R. J.) 142-145 (Adelaide University, 1990).
33. Trombka, T. I. *et al. Astr. J.* **212**, 925 (1977).
34. Fichtel, C. E., Simpson, G. A. & Thompson, D. J. *Astr. J.* **222**, 833 (1978).
35. Gibbs, K. J. *Nucl. Instr. Meth. A* **264**, 67-73 (1988).
36. Gaisser, T. K., Halzen, F., Long, W., Stanev, T. & Zas, E. *Phys. Rev. D* **43**, 314 (1990).
37. Sobel, H. *Nucl. Phys. B (Proc. Suppl.)* **A14**, 125-142 (1990).
38. Heintze, J., Lennert, P., Polenz, S. *et al. Nucl. Phys. B (Proc. Suppl.)* **A14**, 148-152 (1990).
39. Weeks, T. C. *et al. Astrophys. J.* **342**, 379-395 (1989).
40. Akerlof, C. W. *et al. Nucl. Phys. B (Proc. Suppl.)* **A14**, 237-243 (1990).
41. Sommers, P. & Elbert, J. in *Proc. 19th ICRC*, 1985, vol. 3, 457-460.

42. Porter, N. A. & Weekes, T. C. *Mon. Not. R. astr. Soc.* **183**, 205–210 (1978).
43. Akerlof, C. W. *et al.* in *Proc. 21st ICRC, Adelaide*, vol. 2 (ed. Protheroe, R. J.) 135–138 (Adelaide University, 1990).
44. Porter, N. A. & Weekes, T. C. *Astrophys. J.* **212**, 224–226 (1977).
45. Porter, N. A. & Weekes, T. C. *Mon. Not. R. Astr. Soc.* **183**, 205–210 (1978).
46. Porter, N. A. & Weekes, T. C. *Nature* **277**, 199 (1979).
47. Nolan, K., Porter, N. A., Fegan, D. J., Chantell, M. & Weekes, T. C. *Proc. 21st ICRC, Adelaide*, vol. 2 (ed. Protheroe, R. J.) 150 (Adelaide University, 1990).
48. Fegan, D. J., McBreen, B., O'Brien, D. & O'Sullivan, C. *Nature* **271**, 731–732 (1978).
49. Bhat, P. N. *et al.* *Nature* **284**, 433–434 (1984).
50. Porter, N. A. & Weekes, T. C. *Nature* **267**, 500–501 (1977).

51. O'Mongain, E. *Nature* **242**, 136–137 (1973).
52. Huguenin, G. R. & Moore, E. L. *Astrophys. J.* **187**, L57–L58 (1974).
53. Phinney, S. & Taylor, J. H. *Nature* **277**, 117–118 (1979).
54. Rees, M. J. *Nature* **266**, 333–334 (1977).
55. Jelley, J. V., Baird, G. A. & O'Mongain, E. *Nature* **267**, 499–500 (1977).

ACKNOWLEDGEMENTS. We thank C. J. Goebel for giving his time whenever we needed expert help. We thank N. Porter for discussions. This research was supported in part by the University of Wisconsin Research Committee with funds granted by the Wisconsin Alumni Research Foundation, in part by the US Department of Energy, in part by the Xunta de Galicia (Spain) and in part by the Smithsonian Scholarly Studies Research Fund. J.H.M. is a NAS/NRC research associate.

ARTICLES

Crystal structure of insecticidal δ -endotoxin from *Bacillus thuringiensis* at 2.5 Å resolution

Jade Li*, Joe Carroll† & David J. Ellar†

* Medical Research Council Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK

† Biochemistry Department, Cambridge University, Tennis Court Road, Cambridge CB2 1QW, UK

The structure of the δ -endotoxin from *Bacillus thuringiensis* subsp. *tenebrionis* that is specifically toxic to Coleoptera insects (beetle toxin) has been determined at 2.5 Å resolution. It comprises three domains which are, from the N- to C-termini, a seven-helix bundle, a three-sheet domain, and a β sandwich. The core of the molecule encompassing all the domain interfaces is built from conserved sequence segments of the active δ -endotoxins. Therefore the structure represents the general fold of this family of insecticidal proteins. The bundle of long, hydrophobic and amphipathic helices is equipped for pore formation in the insect membrane, and regions of the three-sheet domain are probably responsible for receptor binding.

THE δ -endotoxins are a family of insecticidal proteins produced by *Bacillus thuringiensis* (B.t.) during sporulation, having relative molecular masses (M_r) 60,000–70,000 (60K–70K) in the active form and specific toxicities against insects in the orders of Lepidoptera, Diptera and Coleoptera^{1,2}. These toxins have been formulated into commercial insecticides for three decades³, and now insect-resistant plants are engineered by transformation with Lepidoptera-specific toxin genes^{4–6}. In the bacterium δ -endotoxins are synthesized as protoxins of M_r s 70K–135K and crystallize as a parasporal inclusion $\sim 1 \mu$ in size, in which form they are ingested by the susceptible insect. The microcrystal dissolves in the alkaline pH of the midgut and the protoxin is cleaved by gut proteases to release the active toxin. δ -Endotoxins activated *in vitro* bind specifically and with high affinity ($k_D \approx 0.1$ – 20 nM) to protein receptors on brush-border membrane vesicles derived from the gut epithelium of target insects^{7–9} and create leakage channels of 10–20 Å diameter in the cell membrane¹⁰. *In vivo* such membrane lesions lead to swelling and lysis of the gut epithelium¹¹ and death of the insect ensues through starvation and septicaemia. Active δ -endotoxins of different specificities show five strongly conserved regions in their amino-acid sequences^{1,12}. Exchanging sequence segments in the divergent regions between toxins of different specificities can produce active hybrids showing altered target specificity^{13–15}. We have determined the atomic structure of a

Coleoptera-specific δ -endotoxin (CryIIIA, beetle toxin) from *B.t.* subsp. *tenebrionis*^{16–18} to elucidate the structural basis for target specificity and membrane perforation by this family of proteins.

Structure determination

Parasporal crystals of the beetle toxin contain the full-length 644-residue protoxin¹⁷ as the minor component, and a product of bacterial processing with 57 residues removed from the N-terminus as the major component¹⁹. The latter (M_r 67K) is similar in sequence to the active form of other δ -endotoxins. After solubilization, papain cleavage converts the mixture to the 67K toxin (see legend to Table 1). This was recrystallized in the original crystal form of the parasporal crystals, space group C22₁, and cell dimensions 117.1 by 134.2 by 104.5 Å, containing one molecule per asymmetric unit and 55% solvent by volume¹⁸.

Initial evaluation of derivatives was carried out at 4.5 Å resolution with data collected on the FAST TV diffractometer²⁰ using CuK α radiation. Complete datasets (Table 1) were then collected to 2.5 Å resolution from native crystals using the imaging plate systems at the EMBL outstation at DESY and from the mercury and platinum derivatives on film at SRS Daresbury. The electron density map (Fig. 1) at 2.5 Å resolution calculated with phases from multiple isomorphous replacement (mean figure of merit, 0.63) was easily interpretable and was improved by solvent flattening^{21,22}. A continuous polypeptide chain from residue 61 to residue 644 at the C terminus was traced unambiguously, and most side-chain atoms could be located in the map. The atomic model was built using the graphics program O (ref. 23) and had an initial *R*-factor of 37% for all data to 2.5 Å. After preliminary refinement using the program X-PLOR (ref. 24), the current model, containing 584 amino acid residues and 40 bound water molecules, has an *R*-factor of 19.9% and r.m.s. bond length deviation of 0.017 Å.

Description of the structure

Overview. The beetle toxin is a wedge-shaped molecule with a radius of gyration of 58 Å. As shown in Fig. 2a, it comprises three domains. Domain I, from the N terminus of the 67K toxin to residue 290, is a seven-helix bundle in which a central helix is completely surrounded by six outer helices tilted at about +20° to it (Fig. 3b,c). Domain II, from residues 291 to 500, contains three antiparallel β sheets packed around a hydrophobic core with a triangular cross-section (Fig. 4). Domain III, from residues 501 to 644 at the C terminus is a sandwich of two antiparallel β sheets (Fig. 5). Domains I and III make up the

TABLE 1 Data collection and phasing statistics

Data collection						
Data	Method of collection	Number of crystals	Resolution (Å)	Number of measurements	Unique reflections (% completeness)	R_{merge}
Native	image plate	8	2.5	121,767	27,727 (100)	0.108
CH ₃ HgNO ₃	film	7	2.5	103,623	27,767 (100)	0.095
Hg(CH ₃ COO) ₂	film	5	2.5	60,224	25,919 (94.5)	0.103
<i>cis</i> -Pt(NH ₃) ₂ Cl ₂	film	7	2.5	86,629	25,924 (94.5)	0.107
K ₂ OsO ₄	FAST	1	4.5	21,143	4,680 (100)	0.077
HoCl ₃	FAST	1	4.5	20,013	4,701 (100)	0.069
Phasing statistics						
Derivative	Anomalous data	Number of sites	$R_{\text{deriv}}^{\dagger}$	$R_{\text{Cullis}}^{\ddagger}$	Phasing power§ (resolution, Å)	
CH ₃ HgNO ₃	no	3	0.183	0.715	1.56 (2.5)	
Hg(CH ₃ COO) ₂	yes	6	0.247	0.609	2.28 (2.5)	
<i>cis</i> -Pt(NH ₃) ₂ Cl ₂	no	5	0.185	0.682	1.54 (2.5)	
K ₂ OsO ₄	no	4	0.149	0.757	1.26 (5.5)	
HoCl ₃	no	3	0.095	0.741	1.35 (5.0)	

Protein preparation: Solubilized parasporal crystals from *B.t. subsp. tenebrionis* were incubated at 0.5 mg ml⁻¹ protein with 0.125 units per ml of Agarose-linked papain (Boehringer) in 3.3 M NaBr, 0.05 M sodium phosphate, pH 7.0, and 0.1 mg ml⁻¹ phenylmethylsulphonylfluoride (PMSF) for 30 min at 20 °C. Digestion was stopped by adding tosyl lysinechloromethylketone (TLCK) to 0.125 mg ml⁻¹ and Na₂CO₃ to one fifth volume and removing the enzyme-beads. The 67K beetle toxin was then purified by gel filtration on Sephadex G75 equilibrated with 0.1 M NaHCO₃, pH 10.5, 0.5 M NaBr. **Crystallization:** Single crystals were obtained by microdialysis at a protein concentration of 2.5 mg ml⁻¹ against 0.1 M NaHCO₃, pH 9.5, 1.2 M NaBr at 4 °C overnight, then against 0.1 M NaHCO₃, pH 9.2, 0.5 M NaBr at 16 °C; 3 mM Na₂S₂O₃, 0.1 mM PMSF and 0.1 mg ml⁻¹ TLCK were present in all buffers. Crystals were transferred by stages to 0.05 M 2-(*N*-morpholino)ethanesulphonic acid (MES), pH 6.5, for derivative preparation and mounted in 0.03% low-melting agarose in this buffer during data collection. **Data collection:** Image plate and film data were processed using MOSFLM (Imperial College, London) and CCP4 programs (Daresbury, UK). FAST (ref. 20) data were collected and processed with MADNES⁴⁵, and scaled in 3° batches. **Derivatives:** Crystals were soaked respectively in 0.25 mM CH₃HgNO₃ for 3.5 h, in 1 mM Hg(CH₃COO)₂ for 14 h, in freshly prepared 1 mM *cis*-Pt(NH₃)₂Cl₂ for 21 h, in saturated K₂OsO₄ for 35 h, and in 2 mM HoCl₃ for 3 days. **Phase calculation:** Two heavy-atom sites in each derivative were located from difference Patterson functions, except in the case of Hg(CH₃COO)₂ for which 3 sites were located, and the remaining sites were found by cross-phased difference Fourier. Heavy-atom parameters were refined against centric data and phases calculated for all data using the program PHARE (G. Bricogne). The two low-resolution derivatives were refined against phases calculated from the high-resolution derivatives. Phasing with the three high-resolution derivatives gave an overall figure of merit of 0.61 (25–2.5 Å) and a clearly interpretable map. Including the remaining derivatives slightly improved the connectivity of the map (overall figure of merit 0.63), and four cycles of solvent flattening using a 50% solvent content and a 9 Å radius in mask calculation^{21,22} improved the overall definition of densities. The starting model was built using the program O (ref. 23) with the Bones option for main-chain tracing and the autobuild and manip options for side chains. Refinement by simulated annealing using the program X-PLOR (ref. 24) reduced the *R*-factor from 0.37 to 0.25 without individual *B*-factors, and to 0.23 with restrained individual *B*-factors. The model was adjusted in the loops 154–156, 429–436, and 483–488, and had 40 solvent molecules added, then refined by X-PLOR again. The current model has an *R*-factor of 19.9%, with r.m.s. bond length deviation of 0.017 Å, r.m.s. bond angle deviation of 3.2°, and average atomic *B*-factor of 18 Å².

* $R_{\text{merge}} = \sum_i |I_i - \langle I \rangle| / \sum_i \langle I \rangle$, where I_i are intensity measurements for a reflection, and $\langle I \rangle$ is the mean intensity for this reflection.

† $R_{\text{deriv}} = \sum |F_{\text{PH}} - F_{\text{P}}| / \sum |F_{\text{P}}|$, where F_{PH} is the structure factor amplitude of the derivative crystal and F_{P} is that of the native.

‡ $R_{\text{Cullis}} = \sum ||F_{\text{PH}} \pm F_{\text{P}}| - F_{\text{H}}(\text{calc})| / \sum |F_{\text{PH}} - F_{\text{P}}|$, where F_{P} and F_{PH} are defined as for R_{deriv} , and $F_{\text{H}}(\text{calc})$ is the calculated heavy-atom structure factor amplitude summed over centric data only.

§ Phasing power = $\langle F_{\text{H}} \rangle / E$, the r.m.s. heavy-atom structure factor amplitudes divided by the residual lack of closure error.

bulky end of the molecule. Through their contact one of the two β sheets in domain III is almost entirely buried. To our knowledge (see, for example, ref. 25), the packing of helices in domain I and of sheets in domain II are both novel arrangements.

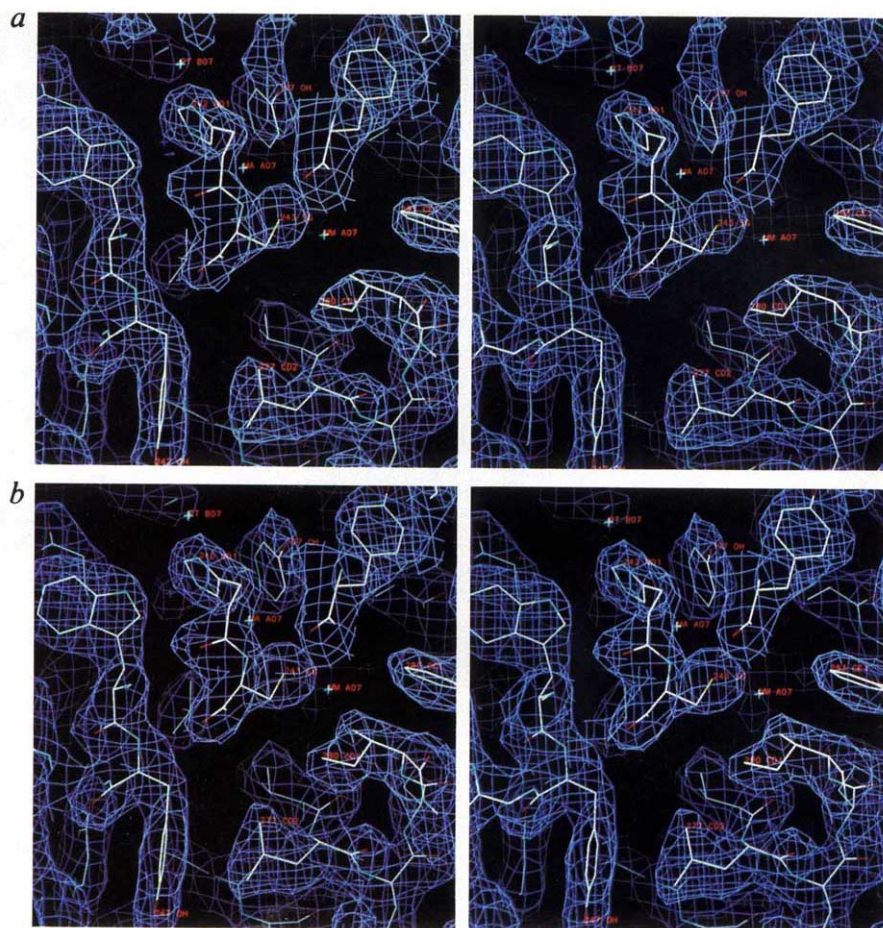
Domain I. The central helix in this seven-helix bundle is α_5 (Fig. 3b,c), which is oriented with its C terminus towards the bulky end of the molecule. Viewed from this end, the outer helices are arranged anticlockwise in the order of α_1 , α_2 , α_3 , α_4 , α_6 and α_7 , with helices α_1 and α_7 adjacent to the β -sheet domains; α_2 is interrupted by a non-helical section and only the leading half, α_{2a} , is packed against α_5 . Figure 3a shows the alignment of amino-acid sequence on the surfaces of the helices. The helices are long, especially α_3 to α_7 , which contain respectively 8, 7, 6, 9 and 7 complete helical turns and hence would be long enough to span the 30-Å thick hydrophobic region of a membrane bilayer. Furthermore, the six outer helices bear a strip of hydrophobic residues (defined by $\Delta G \geq 0$ for transfer from oil to water) down their entire length on the side-facing helix α_5 , so they are amphipathic. In keeping with the general observation that secondary structures are close-packed and bury hydrophobic surfaces²⁶, the helix contact angles in this domain cluster around +20° rather than -50°, giving the bundle a bouquet-like appearance (Fig. 3b). Figure 3c shows the bundle in cross-section. The interhelical space contains 27 aromatic residues which are packed in the edge-to-face fashion²⁷; all polar groups in this region are hydrogen-bonded or in salt bridges.

The concentric arrangement of the seven-helix bundle is distinct from the two-layered type seen in bacteriorhodopsin. There is some resemblance to the pore-forming domain of colicin A²⁸, in which two hydrophobic helices are shielded from solvent by eight amphiphilic helices, but the colicin helices are generally shorter. Like the colicin helices, the bundle in the beetle toxin may be a soluble form of packaging for the hydrophobic and amphiphilic helices that will form pores in the membrane after a large change in conformation.

Domain II. In Fig. 4a and 4b the three sheets of this domain are laid side-by-side, as they would be seen from the solvent. There is an apparent structural duplication between the four-stranded antiparallel sheets, sheet 1 and sheet 2. The chain connections, β_4 , β_3 , β_2 , β_5 and β_8 , β_7 , β_6 , β_9 , respectively, follow the order of +3, -1, -1, +3, which is typical of the 'Greek-key' topology²⁹. From both sheets the inner strands, β_3 and β_2 as well as β_7 and β_6 , extend some 20 Å to the apex of the molecule as two-stranded β ribbons; and at the point of departure from the sheets there is a β -bulge in β_3 and in β_7 to twist the plane of the ribbon by nearly 90° relative to the sheet. The connections between the outer strands cross over the ribbons on the solvent side.

The pseudo-symmetry between these sheets is very approximate. Using the least squares option in O (ref. 23), the sheet region of the strands β_3 and β_2 can be brought to superimpose on that of β_7 and β_6 , with a r.m.s. fit of 0.72 Å for 13 α carbons. But the r.m.s. fit increased to 1.1 Å for 23 α carbons of the

FIG. 1 Electron density map in the neighbourhood of Cys 243, calculated *a*, using combined phases⁴⁶ from multiple isomorphous replacement and solvent flattening, and *b*, using combined experimental and model phases⁴⁶ after refinement by X-PLOR. The refined structure is shown superimposed for reference. Although Cys 243 is a major site of both the methylmercury (MM) and mercuric acetate (MA) derivatives, the methyl mercury site is in a hydrophobic environment compared with the mercuric acetate site.



whole inner strands including the ribbon region, and 1.7 Å for 36 α carbons on all four strands. Nonetheless, the sequence alignment brought by this superposition of the two sheets revealed a low level of internal homology, with seven pairs of equivalent residues (shown in bold) out of 41 aligned α carbons:

```
338 HRIQPHTRFQP(6)SFNYWS(1)NYVSTRPSI(0)GSNDIITSPF(10)NLKFN 395
402 AVANTNLAVWP(0)SAVYSG(1)TKVEFSQYN(3)DEASTQTYDS(7)SWDSI 453
```

The three-stranded sheet 3 is formed by two separate polypeptide segments. The C-terminal segment of domain II contributes the two-stranded ribbon of β_{10} and β_{11} , whereas the N-terminal segment of this domain contributes strand β_1 , which is hydrogen-bonded to β_{11} ; β_1 is followed by a two-turn helix α_8 and an extended chain.

Figure 4c and d shows in side view and in cross-section that the three antiparallel sheets are packed around a triangular hydrophobic core. This brings the strand β_{10} on the edge of sheet 3 into proximity with strand β_4 on the edge of sheet 1, as well as placing the loops at the end of the three β ribbons into a region of about 12 Å radius at the molecular apex. This domain is in contact with helix α_7 of domain I on the face of sheet 3 (Fig. 4c).

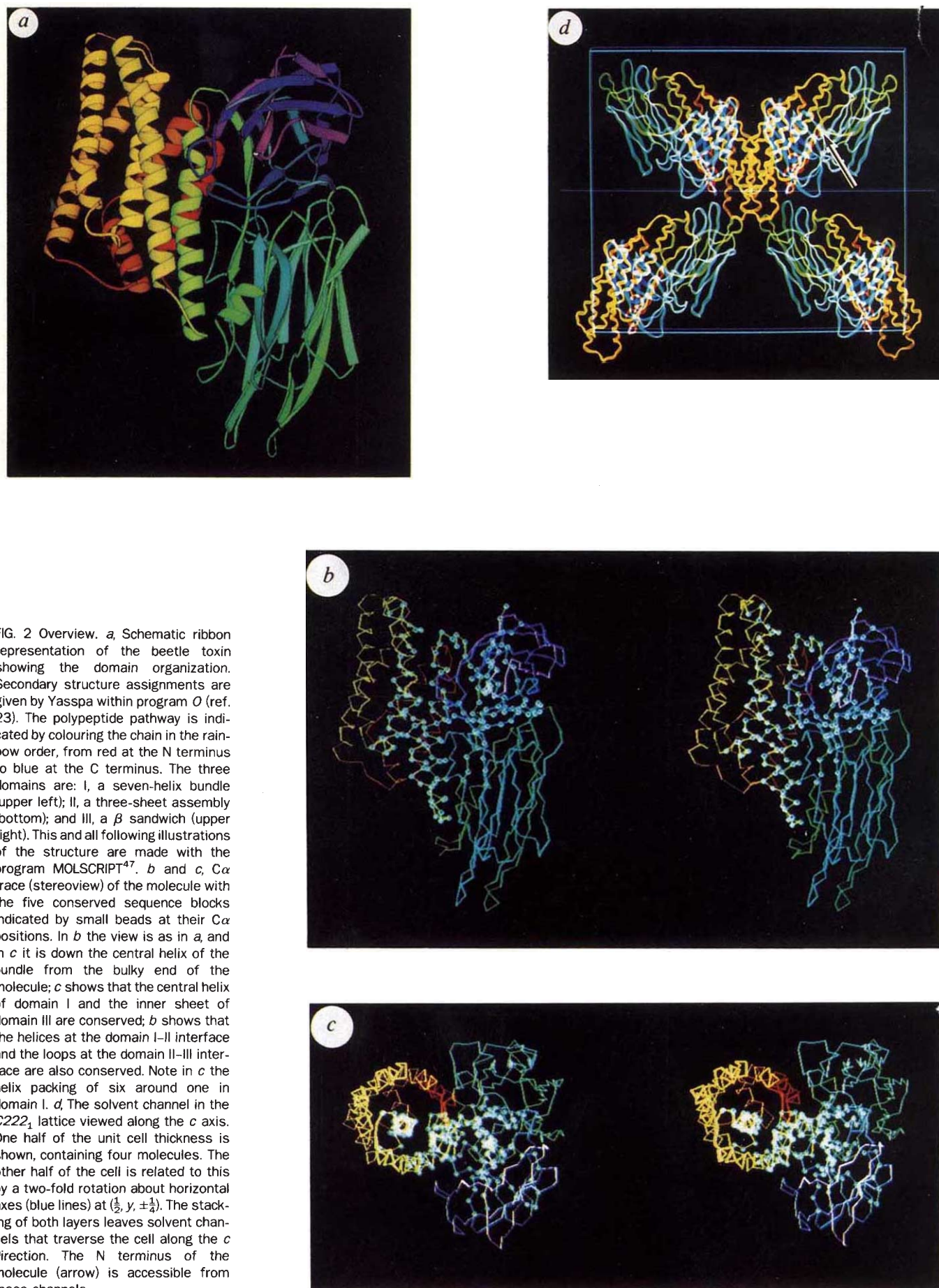
Domain III. Figure 5 is a ribbon drawing of the strands forming the two sheets of the β sandwich. The sheet containing the C-terminal strand is in contact with domain I and will be called the inner sheet. This domain has the 'jelly-roll' topology²⁹, because it can be generated by folding an antiparallel β ribbon which starts with β_{13} (N terminus) and β_{23} (C terminus) on the inner sheet, and ends in the loop between β_{18} and β_{19} on the outer sheet; β_{14} is a short excursion from this ribbon and forms the fifth antiparallel strand of the outer sheet. In addition, small parallel sheets are formed at the edge of the β sandwich through hydrogen bonding of strand β_{12} to β_{16} at the edge of the outer sheet, and β_1 to β_{13} at the edge of the inner sheet.

Distribution of conserved sequences. The core of the beetle toxin molecule encompassing the domain interfaces is built from the five sequence blocks that are highly conserved throughout the δ -endotoxin family¹ (Fig. 2b,c). Block 1, located in the beetle toxin sequence at residues 189–218, corresponds to the central helix (α_5) of the bundle in domain I. Block 2, residues 239–305, overlaps with the latter half of α_6 , and with α_7 and β_1 ; the latter hydrogen-bonds to the edge of the inner sheet in domain III before forming part of the three-stranded sheet 3 in domain II. Block 3, residues 491–538, overlaps with the latter part of β_{11} , where it is hydrogen-bonded to β_1 , and with the loops connecting domains II and III. The remainder of block 3 together with blocks 4 and 5, namely residues 560–569 and 633 to the C terminus, respectively, constitute the three buried strands of the inner antiparallel sheet in domain III. The high degree of conservation of internal residues implies that homologous proteins would adopt a similar fold. Using the beetle toxin structure as a model, we can therefore propose a basis for the insecticidal activity of δ -endotoxins as a family.

Basis of insecticidal function

Solubility. The beetle toxin crystals are isomorphous with the parasporal crystals^{18,19} and show the molecular contacts responsible for solubility behaviour *in vivo*. Four intermolecular salt bridges, Asp 142–Arg 165, Asp 224–Arg 562, Asp 590–Arg 178, and Glu 223–Lys 293, are located at contacts to three different neighbouring molecules. Such salt bridges keep the protoxin crystals insoluble until exposed to the extreme pHs in the insect midgut.

Proteolytic activation. Pro- δ -endotoxins have M_r s of either ~130K or ~70K. Activation by larval gut proteases removes the C-terminal half of the larger protoxins^{30,31} and cleaves them at residue 28 or 29 from the N terminus. The smaller protoxins, such as that of the beetle toxin, are processed only at the N



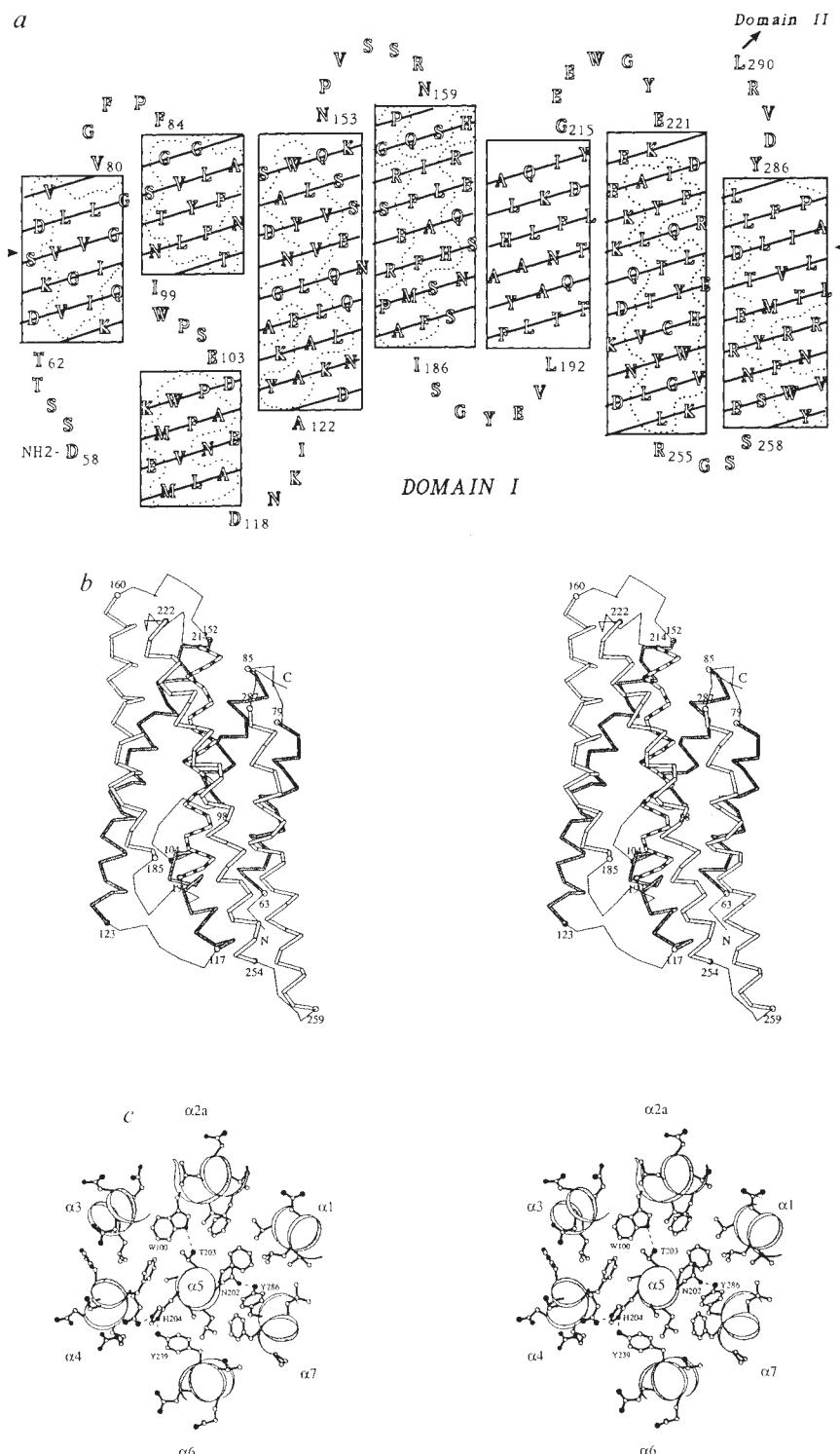


FIG. 3 The seven-helix bundle. *a*, Helical nets showing the position of amino-acid residues along the 7 helices: α_1 (63–79); α_2 (α_{2a} , 85–98 and α_{2b} , 104–117); α_3 (123–152); α_4 (160–185); α_5 (193–214); α_6 (222–254) and α_7 (259–285). The cylindrical surface of the helices are cut longitudinally on the side facing the solvent and flattened to give a view from the interior of the bundle. The top of the drawing corresponds to the bulky end of the whole molecule. Owing to tilting of the outer helices, different helices are in register vertically only at a level indicated by two arrows pointed at α_1 and α_7 ; α_5 is the central helix. Dotted curves outline the strip of hydrophobic residues down the inward surface of the other six helices. *b*, C α trace (stereoview) for the bundle viewed perpendicular to α_5 . The relative tilt of the outer helices to α_5 and that between adjacent outer helices are both about 20°. The C α trace is shaded grey over helices α_1 to α_3 in the back, striped over helix α_5 in the centre, and white over helices α_4 , α_6 , and α_7 in the front. *c*, Cross-section of the bundle at the level indicated by the arrows in *a*, viewed from the bulky end of the molecule. The helical backbone is represented by curly ribbons passing through the C α positions. The outer helices are positioned roughly hexagonally around the central one and tilted relative to it, so the bundle forms a left-handed superhelix. The aromatic side chains are packed in an edge-to-face fashion. Hydrogen bonds are shown for side-chain atoms.

terminus^{19,32} where about 50 residues are removed. The activated δ -endotoxins show a conserved C-terminus, so-called sequence block 5 (ref. 1). Its position as the middle strand of the buried β sheet in domain III precludes further processing from the C terminus. In fact deletion from this site by 4 to 8 residues results in inactive mutants with altered solubility and immunogenicity^{30,33–35}. This is not surprising as the inner sheet can be expected to play a critical part in the structural integrity and stability of the toxins through interaction with the helical bundle.

At the N-terminal cleavage sites the different protoxin sequences show locally similar hydropathy profiles^{36,37}, which would be consistent with a common topology for the N-terminal region of the activated toxins as seen in the helical bundle of

the beetle toxin. In crystals of the beetle toxin, the N terminus at the start of helix α_1 borders on a large solvent channel of about 30 Å diameter that crosses the unit cell along the *c* direction (Fig. 2*d*). This channel could allow access of sporulation-associated proteases to the cleavage site in parasporal crystals¹⁹.

Receptor binding. The insecticidal selectivity of δ -endotoxins is due to high-affinity binding to specific membrane receptors^{7–9,38}, which in three cases seem to be glycoproteins^{38–40}. For several δ -endotoxins the specificity-determining regions have been delimited by exchanging sequence segments between closely related toxins of differing specificities^{13–15}. Guided by the location of secondary structures in the beetle toxin, a plausible alignment of δ -endotoxin sequences was made for the non-

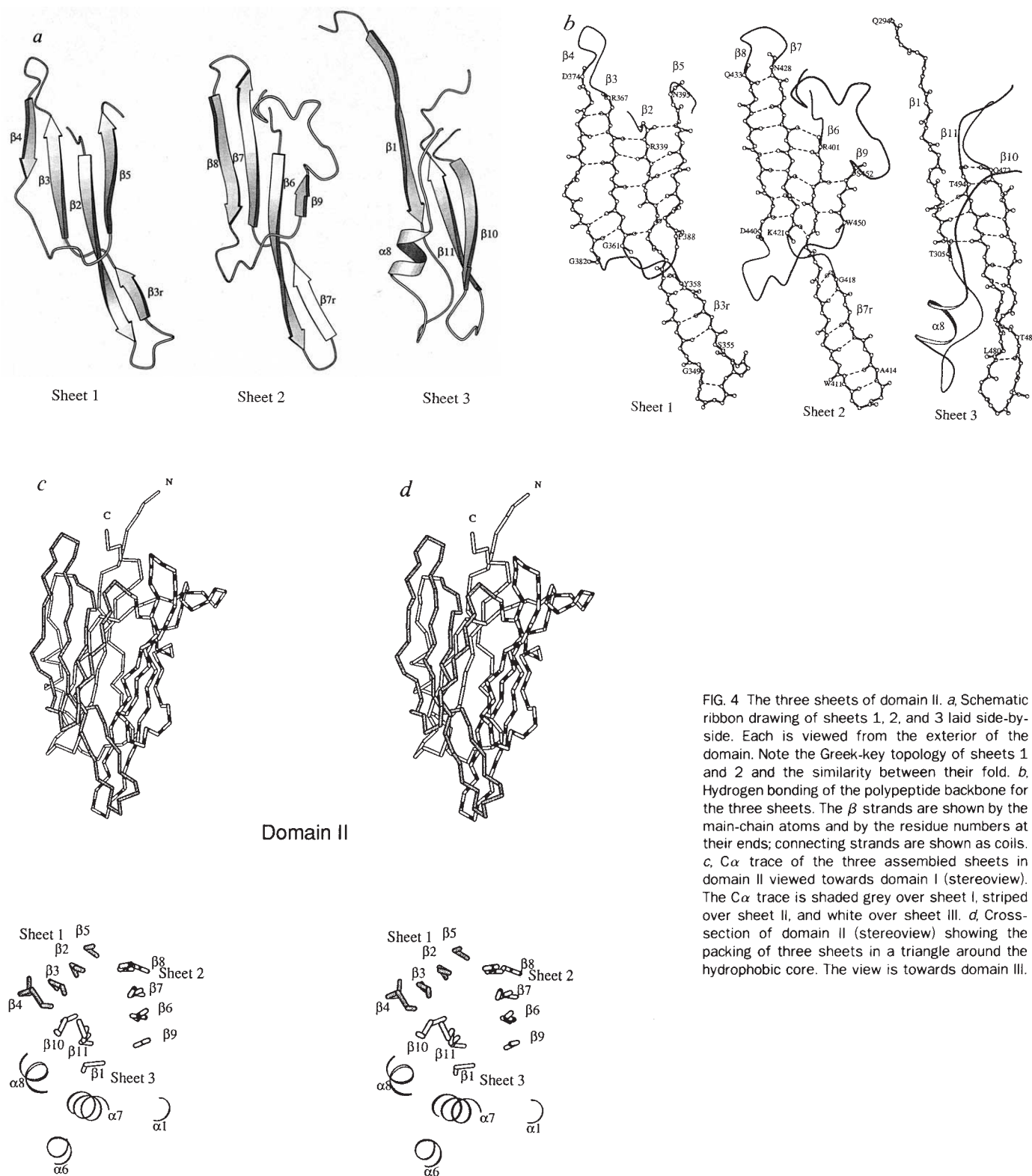


FIG. 4 The three sheets of domain II. *a*, Schematic ribbon drawing of sheets 1, 2, and 3 laid side-by-side. Each is viewed from the exterior of the domain. Note the Greek-key topology of sheets 1 and 2 and the similarity between their fold. *b*, Hydrogen bonding of the polypeptide backbone for the three sheets. The β strands are shown by the main-chain atoms and by the residue numbers at their ends; connecting strands are shown as coils. *c*, $C\alpha$ trace of the three assembled sheets in domain II viewed towards domain I (stereoview). The $C\alpha$ trace is shaded grey over sheet I, striped over sheet II, and white over sheet III. *d*, Cross-section of domain II (stereoview) showing the packing of three sheets in a triangle around the hydrophobic core. The view is towards domain III.

conserved regions (ref. 12, and T. C. Hodgman, unpublished results). Hence the genetically identified specificity-determining regions can be mapped to equivalent positions in the beetle toxin structure, and these fall mainly in domain II. For instance, the dual specificity of CryIIA for Lepidoptera and Diptera, as distinct from the Lepidoptera specificity in the closely related CryIIB, is determined by residues 307–382 of their sequences¹⁴, which corresponds roughly to sheet 1 (Fig. 4a) plus strand β_6 in sheet 2 and the loop leading up to β_7 , whereas the Lepidoptera

specificity of CryIIB is dependent on a longer segment¹⁴ that would include both inner strands of sheet 2. Similarly, the toxicities of CryIA(a) and CryIA(c) to two lepidopteran insects depend on three segments termed x, y and z (ref. 15): amino-acid substitutions in y can reduce toxicity by up to 2,000-fold, and segments x and y interact in determining specificity. Aligned with the beetle toxin structure, segment x corresponds roughly to the outer strands β_4 and β_5 of sheet 1 and the whole of sheet 2, including the loop entering β_{10} in sheet 3; y corresponds to

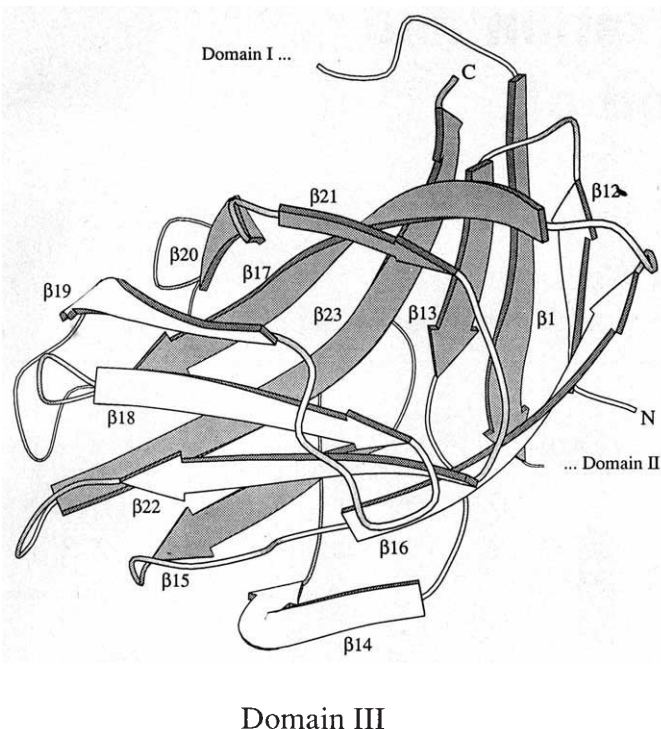


FIG. 5 Domain III, schematic ribbon representation of the β sandwich. β strands forming the inner sheet are shaded grey. The topology of an eight-stranded 'jelly-roll' can be seen by following the β hairpin starting with β_{13} , β_{15} and β_{23} in the inner sheet, continuing to β_{16} and β_{22} in the outer sheet, then β_{17} and β_{21} , β_{20} in the inner sheet, and ending with β_{18} and β_{19} in the outer sheet. β_{14} is an excursion from the hairpin and forms a fifth antiparallel strand of the outer sheet. Small parallel β sheets are added to one edge of the β sandwich, by hydrogen bonding of β_1 to β_{13} in the inner sheet and β_{12} to β_{16} in the outer sheet. Residue numbers in the β strands are: β_{12} , 502–506; β_{13} , 509–513; β_{14} , 519–525; β_{15} , 536–541; β_{16} , 547–554; β_{17} , 558–569; β_{18} , 573–579; β_{19} , 585–591; β_{20} , 604–609; β_{21} , 611–614; β_{22} , 619–625; and β_{23} , 631–643.

strand β_{10} of sheet 3 and the loop connecting β_{10} and β_{11} ; and z extends from β_{11} to the C-terminal activation site. Furthermore, the interaction between x and y can be understood in terms of the proximity between β_4 on the edge of sheet 1 and β_{10} on the

edge of sheet 3. Although z was inferred¹⁵ to extend into domain III, the combined evidence from genetics and receptor-binding assays *in vitro* for Lepidoptera toxins^{9,41} correlates receptor recognition with sequence variations within domain II. We note that the β ribbons from all three sheets terminate in loops in a small region on the molecular apex, in a manner reminiscent of the complementarity-determining region of immunoglobulins.

Pore formation. The common mechanism of epithelial cell disruption by δ -endotoxins of widely different specificities is believed to be the formation of lytic pores of 10 to 20 Å diameter in the insect membrane¹⁰. The structure of the beetle toxin displays an apparatus for pore formation in the long, hydrophobic and amphipathic helices of domain I which could penetrate the membrane. Between the crystal structure in which the bouquet-like helical bundle internalizes all the hydrophobic surfaces, and the unknown pore structure where hydrophobic surfaces would be in intimate contact with the membrane lipids, large conformational changes must occur. In the absence of a full characterization of the pore-forming process, we propose the following by extrapolation from the crystal structure.

The trigger for the conformational changes may be provided by receptor binding and the consequent interaction of toxin with the membrane bilayer. Membrane insertion follows rapidly, so that a major part of the bound δ -endotoxin cannot be displaced from the brush-border vesicles by other toxins recognizing the same receptor sites^{7,9}. As domain II and probably its apical region are most likely to bind the membrane receptors, the helices are expected to insert with the 'domain II end' (see Fig. 2a) oriented towards the cytoplasm. If helical hairpins are to initiate the membrane penetration, as probably happens for colicin^{28,42,43}, they will probably be linked at the domain II end. So either of the helix pairs α_6 – α_7 or α_4 – α_5 could be the likely initiator. The α_6 – α_7 pair is favoured because it forms part of the conserved interface with domain II and is well positioned to sense the receptor binding. On the other hand, helix α_5 is the most conserved throughout the family of δ -endotoxins. Point mutations in α_5 reduce toxicity of a Lepidoptera toxin without reducing binding to membranes⁴⁴. Proteolysis in the interhelical loops at the domain III end, as in the α_3 – α_4 loop^{19,32}, may facilitate release of the helix pairs from the tertiary structure of the bundle. The insertion of a hairpin can create a defect in the membrane, allowing the rest of domain I to participate in pore formation in a cooperative manner. □

Received 22 July; accepted 19 September 1991.

- Höfte, H. & Whiteley, H. R. *Microbiol. Rev.* **53**, 242–255 (1989).
- Ellar, D. J. *et al.* in *Molecular Biology of Microbial Differentiation* (eds Hoch, J. A. & Setlow, P.) 230–240 (Am. Soc. Microbiol., Washington, DC, 1985).
- Wilcox, E. R. *et al.* in *Protein Engineering: Applications in Science, Medicine and Industry* (eds Inoué, M. & Sarma, R.) 395–413 (Academic, New York, 1986).
- Vaeck, M. *et al.* *Nature* **328**, 33–37 (1987).
- Perlak, F. J., Fuchs, R. L., Dean, D. A., McPherson, D. L. & Fischhoff, D. A. *Proc. natn. Acad. Sci. U.S.A.* **88**, 3324–3328 (1991).
- Barton, K. A., Whiteley, H. R. & Yang, N. S. *Plant Physiol.* **85**, 1103–1109 (1987).
- Hofmann, C., Lüthy, P., Hütter, R. & Pliska, V. *Eur. J. Biochem.* **173**, 85–91 (1988).
- Hofmann, C. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **85**, 7844–7848 (1988).
- Van Rie, J., Janssens, S., Höfte, H., Degheele, D. & Van Mellaert, H. *Appl. Envir. Microbiol.* **56**, 1378–1385 (1990).
- Knowles, B. H. & Ellar, D. J. *Biochim. biophys. Acta* **924**, 509–518 (1987).
- Endo, Y. & Nishitsutsuji-Uwo, J. *J. Invertebr. Path.* **36**, 90–103 (1980).
- Hodgman, T. C. & Ellar, D. J. *J. DNA Sequ. Map.* **1**, 97–106 (1990).
- Ge, A. Z., Shivarova, N. I. & Dean, D. H. *Proc. natn. Acad. Sci. U.S.A.* **86**, 4037–4041 (1989).
- Widner, W. R. & Whiteley, H. R. *J. Bact.* **172**, 2826–2832 (1990).
- Schnepf, H. E., Tomczak, K., Ortega, J. P. & Whiteley, H. R. *J. Biol. Chem.* **265**, 20923–20939 (1990).
- Krieg, A., Hüger, A. M., Langenbruch, G. A. & Schnetter, W. *J. appl. Entomol.* **96**, 500–508 (1983).
- Höfte, H., Seurinck, J., Van Houtven, A. & Vaeck, M. *Nucleic Acids Res.* **15**, 7183 (1987).
- Li, J., Henderson, R., Carroll, J. & Ellar, D. J. *J. molec. Biol.* **199**, 543–545 (1988).
- Carroll, J., Li, J. & Ellar, D. J. *Biochem. J.* **261**, 99–105 (1989).
- Arndt, U. W. *Meth. Enzym.* **114**, 472–485 (1985).
- Wang, B. C. *Meth. Enzym.* **115**, 90–112 (1985).
- Leslie, A. G. W. *Acta crystallogr.* **A43**, 134–136 (1987).
- Jones, T. A., Zou, J.-Y., Cowan, S. W. & Kjeldgaard, N. *Acta crystallogr.* **A47**, 110–119 (1991).
- Brünger, A. T. *J. molec. Biol.* **203**, 803–816 (1988).
- Chothia, C. & Finkelstein, A. V. *Rev. Biochem.* **59**, 1007–1039 (1990).

- Chothia, C. *J. molec. Biol.* **105**, 1–15 (1976).
- Burley, S. K. & Petsko, G. A. *Science* **229**, 23–28 (1985).
- Parker, M. W., Pattus, F., Tucker, A. D. & Tsernoglou, D. *Nature* **337**, 93–96 (1989).
- Richardson, J. S. *Adv. Prot. Chem.* **34**, 167–339 (1981).
- Höfte, H. *et al.* *Eur. J. Biochem.* **161**, 273–280 (1986).
- Choma, C. T. & Kaplan, H. *Biochemistry* **29**, 10971–10977 (1990).
- Nicholls, C. N., Ahmad, W. & Ellar, D. J. *J. Bact.* **171**, 5141–5147 (1989).
- MacIntosh, S. C., McPherson, S. L., Perlak, F. J., Marrone, P. G. & Fuchs, R. L. *Biochem. biophys. Res. Commun.* **170**, 665–672 (1990).
- Schnepf, H. E. & Whiteley, H. R. *J. Biol. Chem.* **260**, 6273–6280 (1985).
- Adang, M. J. *et al.* *Gene* **36**, 289–300 (1985).
- Sekar, R., Thompson, D. V., Maroney, M. J., Bookland, R. G. & Adang, M. J. *Proc. natn. Acad. Sci. U.S.A.* **84**, 7036–7040 (1987).
- Chungjatupornchai, W., Höfte, H., Seurinck, J., Angsuthanasombat, C. & Vaeck, M. *Eur. J. Biochem.* **173**, 9–16 (1988).
- Knowles, B. H., Thomas, W. E. & Ellar, D. J. *FEBS Lett.* **168**, 197–202 (1984).
- Knowles, B. H. & Ellar, D. J. *J. Cell Sci.* **83**, 89–101 (1986).
- Haider, M. Z. & Ellar, D. J. *Biochem. J.* **248**, 197–201 (1987).
- Visser, B., Munsterman, E., Stoker, A. & Dirse, W. G. *J. Bact.* **172**, 6783–6788 (1990).
- Lakey, J. H., Baty, D. & Pattus, F. *J. molec. Biol.* **218**, 639–653 (1991).
- Song, H. Y., Cohen, F. S. & Cramer, W. A. *J. Bact.* **173**, 2927–2934 (1991).
- Ahmad, W. & Ellar, D. J. *FEMS Microbiol. Lett.* **68**, 97–104 (1990).
- Messerschmidt, A. & Pfugrath, J. W. *J. appl. Crystallogr.* **20**, 306–315 (1987).
- Read, R. J. *Acta crystallogr.* **A42**, 140–149 (1986).
- Kraulis, P. J. *J. appl. Crystallogr.* (in the press).

ACKNOWLEDGEMENTS. We thank P. R. Evans, A. G. W. Leslie and R. Henderson for advice and encouragement; K. Wilson and Z. Dauter for help with the image plate system; K. Nagai and P. J. McLaughlin for help in collecting film data; T. A. Jones for advice on model building; SERC Daresbury Laboratory and EMBL Outstation at DESY for use of synchrotron facilities; and T. Woollard and K. Hopkins for maintaining the rotating anodes. D.J.E. and J.C. acknowledge the support of the AFRC.

Stress-induced oligomerization and chromosomal relocation of heat-shock factor

J. Timothy Westwood, Joachim Clos & Carl Wu

Laboratory of Biochemistry, National Cancer Institute, National Institutes of Health, Bg 37, Rm 4C-09, Bethesda, Maryland 20892, USA

The induction of heat-shock transcription factor (HSF) binding to DNA is accomplished by a heat-induced oligomerization. The transition to the induced state is accompanied by a chromosomal redistribution of HSF to the heat-shock puff sites. Over 150 additional chromosomal sites also accumulate HSF, including developmental loci that are repressed during heat shock. These findings suggest an unforeseen role for HSF as a repressor of normal gene activity during heat stress.

ORGANISMS respond to a rise in temperature by rapidly increasing the expression of heat-shock proteins (see ref. 1 for review). Since the discovery of the chromosomal response to heat shock², elucidation of the regulation and function of heat-shock proteins at a molecular level has progressed considerably. Heat-shock proteins are now known to assist protein folding, protein-protein interactions, and the translocation of proteins across cellular compartments^{3,4}. The heat-shock response is also used as a model system for the analysis of inducible genes in prokaryotes and eukaryotes. In eukaryotes, transcriptional control of heat-shock genes is mediated through a conserved DNA sequence, the heat-shock element (HSE), composed of three repeats of a five-base module (nGAAn) arranged in inverted orientation⁵⁻⁷. The HSE is the site of interaction with a transcriptional activator, heat-shock factor, HSF (for reviews see refs 8, 9). Complementary DNAs for HSF have been cloned from yeast, fruitfly, tomato, human and mouse¹⁰⁻¹⁷. Despite widespread conservation of heat-shock proteins and the HSE sequence in eukaryotes, the predicted amino-acid sequences of the cloned HSFs are highly divergent, except for the DNA-binding domain and several heptad repeats of hydrophobic residues, or leucine-zipper motifs.

The regulation of HSF activity occurs at several levels, which differ between higher and lower eukaryotes. In the yeasts *Saccharomyces cerevisiae*^{10,11} and *Kluyveromyces lactis*¹⁴, HSF binding to DNA is constitutive and independent of the environmental temperature; heat shock results in an increase in phosphorylation of the *S. cerevisiae* factor that is associated with transcriptional activation¹⁰. In fruitflies, vertebrates and plants, binding of HSF to DNA is low or undetectable in unshocked cells, and a heat-shock stimulus is required to convert HSF to a form that binds with high affinity to the HSE (refs 8, 13). In addition, there is evidence for a heat-induced phosphorylation of HSF in mammalian cells¹⁸.

With respect to the induction of HSF binding to DNA in higher eukaryotes, the inactive HSF proteins seem to be able to respond directly or indirectly to changes in the physical or chemical environment. Thus, HSF binding can be induced *in vitro* by treatment of crude or fractionated cytoplasmic extracts with heat¹⁸, detergent and low pH¹⁹, an oxidizing agent²⁰ and, in the case of *Drosophila* HSF, by reaction with antisera raised to the active form of the protein²¹. The inducible activity of

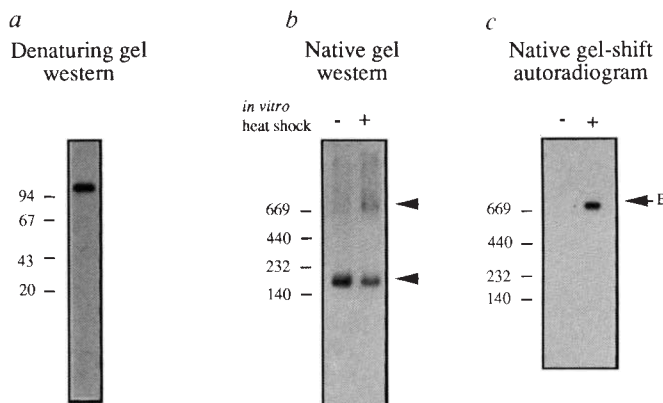


FIG. 1 Native and denatured states of HSF. *a*, Western blot analysis of S2 cell cytoplasmic proteins after SDS-PAGE. *b*, Western blot analysis of unshocked (–) and *in vitro* heat-shocked (+) cytoplasmic proteins after pore exclusion limit electrophoresis on a nondenaturing gel. *c*, Autoradiogram showing ³²P-labelled HSE:HSF complexes formed with unshocked and heat-shocked cytoplasmic proteins and separated as in *b*.

METHODS. Cytoplasmic extracts (24 mg protein ml⁻¹) were prepared at 4 °C essentially according to Dignam *et al.*⁵³. Briefly, 200 ml unshocked S2 cells (~1 × 10⁹ cells) were pelleted and homogenized in two volumes of buffer A, the homogenate was centrifuged for 20 min at 25,000g, and the supernatant was recentrifuged for 60 min at 100,000g. The 100,000g supernatant was the cytoplasmic extract. *In vitro* heat shock of the extract was performed by incubation at 34 °C for 10 min. Aliquots (2.5 μl and 1.5 μl) of extract were analysed by SDS-PAGE and pore exclusion limit electrophoresis, respectively. SDS-PAGE (10% polyacrylamide) was done according to standard procedures and proteins were electrophoretically transferred onto a PVDF (polyvinylidene difluoride) membrane in 48 mM Tris, 39 mM glycine (transfer buffer) using the semidry technique. Pore exclusion limit electrophoresis on nondenaturing gel was done essentially as described previously¹² except that the polyacrylamide gradient was 4–24% and the gel dimensions were decreased to 9 × 10 × 0.15 cm. For transferring native proteins, the gel was incubated for 10 min at 70 °C in transfer buffer containing 0.25% SDS before electroblotting in transfer buffer at 4 °C. Membranes were stained with a 0.2% Ponceau S, 3% sulphosalicylic acid, 3% TCA to visualize *M_r* markers (Pharmacia). Membranes were processed with primary antibody (1:1,000 dilution of rabbit α-HSF) and secondary antibody (1:40,000 dilution of goat anti-rabbit IgG linked to alkaline phosphatase), developed with a luminescent substrate according to the manufacturer's protocol (Tropix), and exposed to Kodak XAR-5 film. Specificity of the α-HSF serum was confirmed by the lack of staining of the 110K band using either preimmune serum or immune serum with excess cloned HSF protein. Pore exclusion limit electrophoresis of HSE:HSF complexes was done by incubation of ³²P-labelled HSE with unshocked and *in vitro* heat-treated cytoplasmic extracts. Extract (1 μl) was incubated for 15 min at 4 °C with 3 μl stock solution of nucleic acids and BSA such that the final reaction contained 250 ng *E. coli* DNA, 50 ng poly d(N)5, 1.25 μg poly(dI.dC)-poly(dI.dC), 100 μg BSA, and 1 fmol of a ³²P-labelled HSE²⁵. Samples were analysed by nondenaturing gel electrophoresis as described above. The gels were fixed, stained with Coomassie blue to visualize the markers, dried and autoradiographed. Polyclonal antibodies to HSF were prepared by injecting rabbits with 500 μg cloned *Drosophila* HSF purified from *E. coli* extracts¹² followed by two injections of 250 μg each at 3-week intervals.

HSF in *Drosophila* cells is unregulated when it is expressed in *Escherichia coli*; the bacterially produced HSF is active as a DNA-binding transcription factor even when expressed at a low temperature¹². The intrinsic activity of HSF is therefore apparently subject to negative control when the protein is translated under normal conditions in *Drosophila* cells. Both cloned and naturally activated *Drosophila* HSF proteins self-associate as trimers or hexamers^{12,22}. This unusual degree of multimerization for a transcription factor has led to the suggestion that HSF binding is regulated by means of a change in the state of oligomerization on heat shock.

Here we report the properties of *Drosophila* HSF under normal and heat-shock conditions. Under normal conditions, HSF appears in the nucleus as a small, inactive oligomer which is converted on heat shock to an active multimer. Indirect immunofluorescent staining of salivary gland polytene chromosomes reveals that the oligomeric transition is accompanied by a relocalization of the transcription factor from a general chromosomal distribution to the heat-shock puff sites. Unexpectedly, many other chromosomal loci also stain positively for HSF, suggesting a role for HSF in the repression of cellular transcription during heat stress.

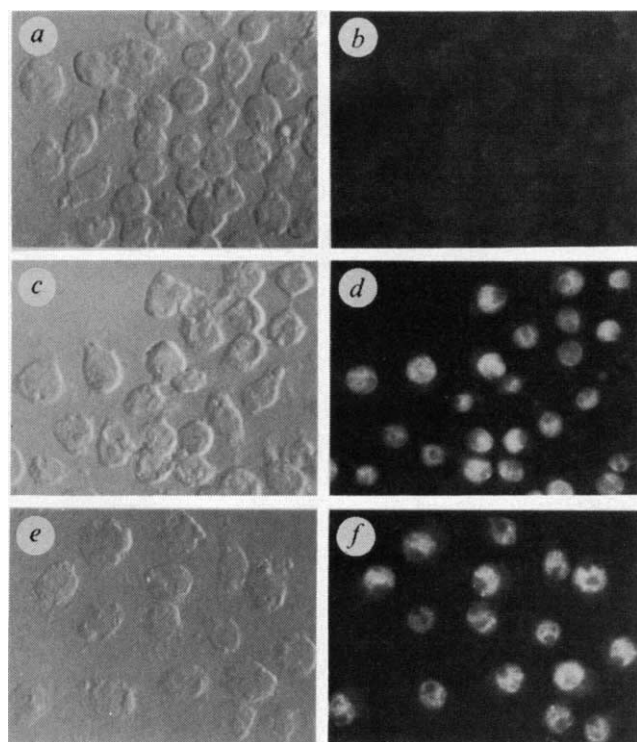


FIG. 2 Nuclear localization of HSF. *a, c, e*, *Drosophila* S2 cells visualized by differential interference microscopy. *b, d, f*, S2 cells stained with preimmune serum (*b*), or rabbit α -HSF serum (*d, f*), followed by FITC-conjugated secondary antibody. Cells were unshocked in *a, b, c, d* and heat shocked in *e, f*. METHODS. S2 cells were grown at 22 °C in plastic chamber slides (Nunc) and heat shocked for 15 min at 37 °C. Cells were fixed in 4% paraformaldehyde in PBS (4 °C) (pH was adjusted to 7.2) for 1–2 min. After replacement with fresh fixative, fixation continued for 15 min at room temperature followed by three washes with PBS. An alternative procedure using cold methanol as fixative produced similar results. The primary antibody was a 1:1,000 dilution of rabbit α -HSF or preimmune serum in PBS, 0.5% BSA, 0.5% Triton X-100. After incubation for 60 min, cells were rinsed thrice with PBS and washed with 0.5% BSA in PBS for 30 min. The secondary antibody was a 1:100 dilution of FITC-conjugated, goat anti-rabbit F(ab')₂ (Cappel/Organon-Technica) in 0.5% BSA, 0.5% Triton X-100, PBS. After incubation for 60 min, cells were washed as above, and mounted with 70% glycerol in PBS. Fluorescent photomicrographs were taken under ultraviolet epi-illumination with a $\times 40$ objective (Nikon) and Fujichrome 400 film (20 s exposure).

Structural alteration on heat shock

HSF proteins have sequence conservation at two adjacent regions near the amino terminus that are required for high-affinity binding to the HSE. For *Drosophila* and *S. cerevisiae* HSFs, deletion of a region bearing conserved leucine zipper motifs abolished the multimeric association and high-affinity binding to the HSE^{11,12,23}. These findings suggested that the affinity of wild-type *Drosophila* HSF for DNA could be regulated by changing the oligomerization state. We tested this hypothesis by analysing the native molecular mass of the active and inactive forms of HSF, using western blotting after pore exclusion limit electrophoresis on nondenaturing gels²⁴. Such an analysis of cytoplasmic proteins from unshocked *Drosophila* tissue culture cells (Schneider line 2, S2) revealed a single complex migrating at a native relative molecular mass of about 220,000 (M_r , 220K) that reacted with polyclonal antibodies raised to bacterially produced HSF protein (α -HSF) (Fig. 1*b*, lane 1). The specificity

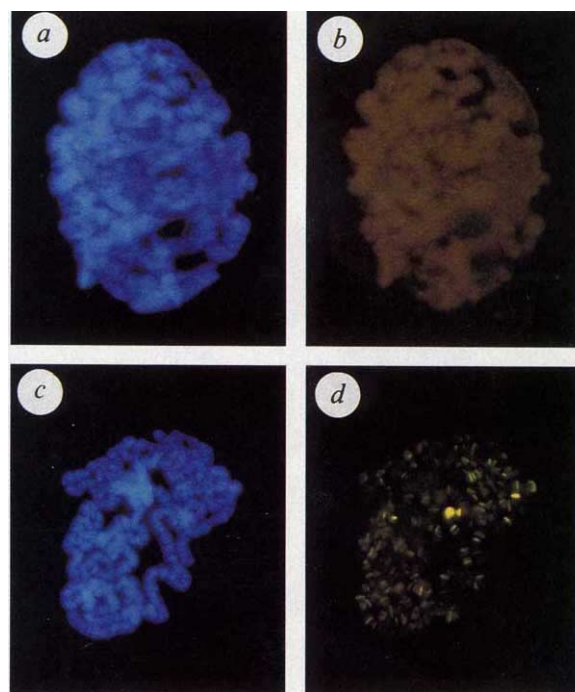


FIG. 3 Localization of DNA and HSF in polytene nuclei from unshocked (*a, b*) and heat-shocked (*c, d*) salivary gland squashes. DNA was stained with bisbenzimidazole (Hoechst 33342) (*a, c*) and HSF with polyclonal rabbit α -HSF antibodies (*b, d*). Some unshocked nuclei show specific staining over the general diffuse pattern at one heat-shock locus (63BC). 63BC is the site for *hsp83*, which carries the highest affinity HSE sequences and is the earliest to respond to heat stress. It is probable that this staining pattern represents the early effects of stress induced during dissection of the salivary glands.

METHODS. Late third instar larvae grown at room temperature (23–24 °C), and larvae heat shocked for 25 min at 36.5 °C were quickly dissected in PBS. Salivary glands were fixed for 3–5 min in 3.7% formaldehyde, 50% acetic acid, squashed on microscope slides and stored in 95% ethanol as described^{27–29}. Similar results were obtained with an additional prefixing step (3.7% formaldehyde, 1% Triton X-100, in PBS for 30 s) (G. Yim and C.W. unpublished results). Slides were washed twice for 30 min in PBS, incubated with 0.5% BSA, 0.1% Tween 20 in PBS (BTP) for 30 min, incubated with a 1:1,000 dilution of α -HSF in BTP, washed twice with PBS, 0.1% Tween 20 for 10 min, incubated with a 1:100 dilution of FITC-conjugated goat anti-rabbit, affinity-purified F(ab')₂ fragments (Cappel), washed with PBS, 0.01% Tween 20 for 10 min, stained for 5 min with 1 μ g ml⁻¹ Hoechst 33342 in PBS, and washed finally for 10 min in PBS, 0.01% Tween 20. Slides were mounted in 1 mg ml⁻¹ phenylenediamine, 70% glycerol in PBS, and analysed by fluorescence microscopy using a Nikon Planapo $\times 40$ objective, Kodak Ektachrome 200 film and exposures of 5 s for DNA (blue) and 40 s (*b*) and 20 s (*d*) for HSF (yellow).

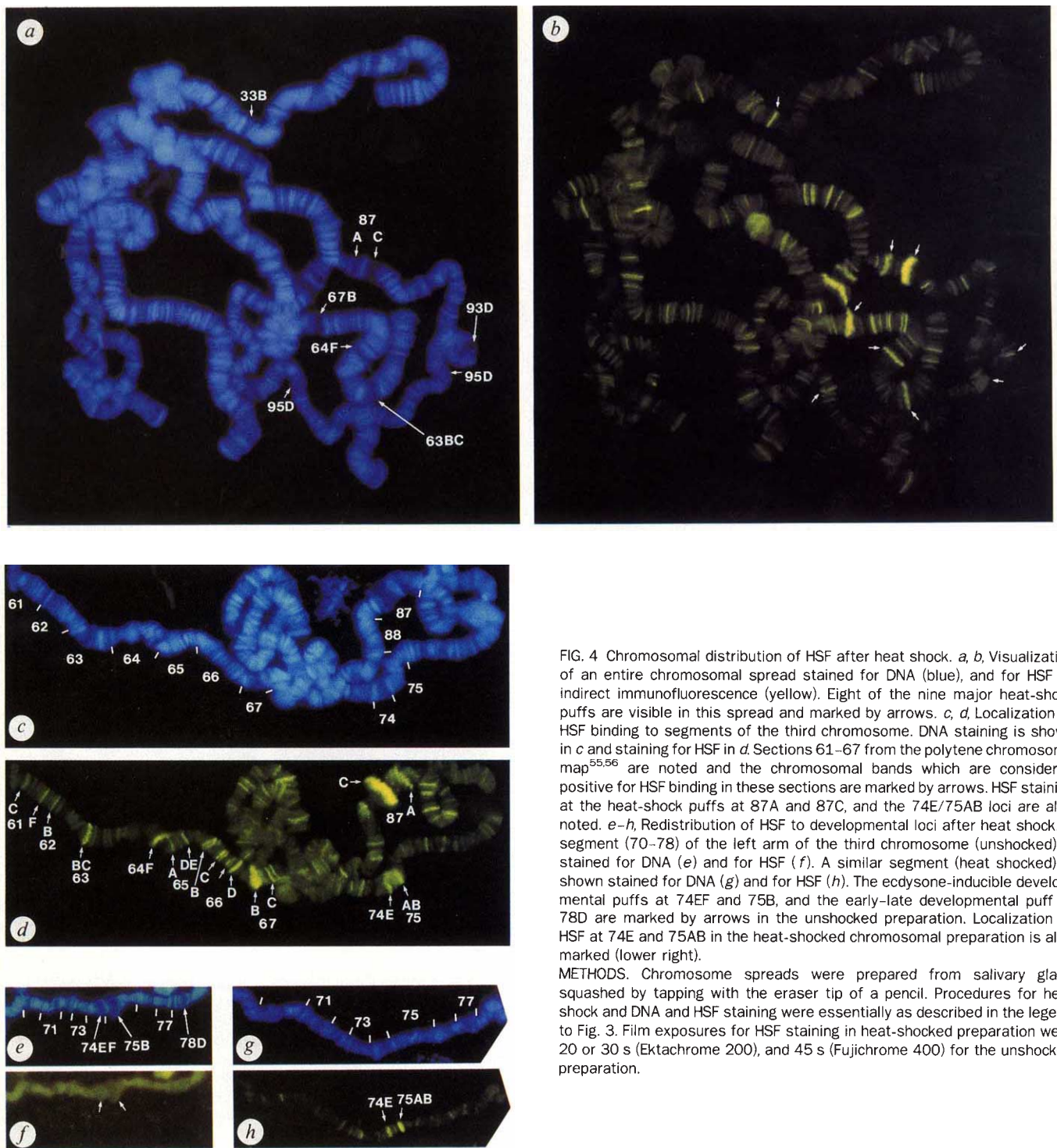


FIG. 4 Chromosomal distribution of HSF after heat shock. *a, b*, Visualization of an entire chromosomal spread stained for DNA (blue), and for HSF by indirect immunofluorescence (yellow). Eight of the nine major heat-shock puffs are visible in this spread and marked by arrows. *c, d*, Localization of HSF binding to segments of the third chromosome. DNA staining is shown in *c* and staining for HSF in *d*. Sections 61–67 from the polytene chromosome map^{55,56} are noted and the chromosomal bands which are considered positive for HSF binding in these sections are marked by arrows. HSF staining at the heat-shock puffs at 87A and 87C, and the 74E/75AB loci are also noted. *e–h*, Redistribution of HSF to developmental loci after heat shock. A segment (70–78) of the left arm of the third chromosome (unshocked) is stained for DNA (*e*) and for HSF (*f*). A similar segment (heat shocked) is shown stained for DNA (*g*) and for HSF (*h*). The ecdysone-inducible developmental puffs at 74EF and 75B, and the early-late developmental puff at 78D are marked by arrows in the unshocked preparation. Localization of HSF at 74E and 75AB in the heat-shocked chromosomal preparation is also marked (lower right).

METHODS. Chromosome spreads were prepared from salivary gland squashed by tapping with the eraser tip of a pencil. Procedures for heat shock and DNA and HSF staining were essentially as described in the legend to Fig. 3. Film exposures for HSF staining in heat-shocked preparation were 20 or 30 s (Ektachrome 200), and 45 s (Fujichrome 400) for the unshocked preparation.

of the α -HSF antibody on the western blot was demonstrated by unique staining of the 110K HSF subunit²⁵ among total cytoplasmic proteins separated by SDS-PAGE (Fig. 1*a*). *In vitro* heat treatment of the cytoplasmic extract resulted in a partial depletion of the 220K species and the appearance of a 690K HSF species, in addition to a small amount of smeared, higher M_r material (Fig. 1*b*, lane 2). When this sample containing both small and large HSF complexes was analysed for DNA-binding activity by the mobility shift technique using pore exclusion limit electrophoresis, the 690K HSF species, but not the 220K species, showed specific binding to ³²P-labelled HSE (Fig. 1*c*, lane 2). Hence, the 220K HSF species represents the inactive form of HSF, which is convertible to the active 690K species on heat treatment.

The large change in the native state of HSF measured above strongly suggests a change in the degree of HSF oligomerization on heat shock. Assuming a globular conformation, the results indicate that the inactive form of HSF could be composed of a homodimer of the 110K subunit, or a heteromer of one HSF subunit and one or more unidentified (inhibitory) proteins. Alternatively, the 220K species could simply reflect the apparent mobility of a highly asymmetric HSF monomer. Additional gel filtration and sedimentation analyses will be required to establish the exact oligomeric state of inactive HSF.

Nuclear localization before heat shock

The cytological location of *Drosophila* HSF was determined in S2 cells by indirect immunofluorescence analysis. As expected

for an active DNA-binding protein, HSF was localized predominantly in cell nuclei after heat shock (Fig. 2e, f). The pattern of immunostaining was over the entire nucleus with areas of punctate staining, with a distinct lack of staining in the nucleolus. In unshocked cells, HSF staining was also localized in nuclei, the staining being more uniform over the entire nucleus, including weak staining in some nucleolar regions (Fig. 2c, d). Nuclear staining of HSF in unshocked cells was confirmed in six independent experiments. No staining was observed with preimmune serum (Fig. 2a, b), nor with α -HSF antibody in the presence of excess HSF protein (data not shown). These results indicate that translocation of HSF to the cell nucleus occurs before any heat-shock stimulus, consistent with the extremely rapid response to the stress signal.

The finding that HSF is localized in the nuclei of unshocked cells was surprising, in view of previous biochemical studies which placed the (inactive) HSF protein in the cytoplasmic fraction of *Drosophila* and mammalian cells^{18,19,21}. To confirm that the S2 cells grown at 24–25 °C for immunostaining were not heat shocked inadvertently, we determined the distribution of HSF in isolated nuclear and cytoplasmic fractions of S2 cells grown under identical conditions. HSF was found in the cytoplasmic fraction of unshocked S2 cells and in the nuclear fraction of heat-shocked cells (data not shown). The apparent discrepancy between the nuclear staining of HSF and its detection in the subcellular cytoplasmic fraction is not difficult to reconcile. In all likelihood, HSF is localized in the nucleus under nonshock conditions, but its nuclear affinity may be insufficient to prevent leakage into the cytoplasmic fraction during cell disruption. We confirmed this possibility by immunostaining nuclei isolated after homogenization of unshocked and heat-shocked S2 cells. There was a significant decrease in the amount of HSF staining in the nuclear fraction prepared from unshocked cells, whereas nuclei from heat-shocked cells retained a high level of HSF, until extracted with 0.4 M NaCl (data not shown). A similar behaviour of the oestrogen receptor has been reported^{26,27}.

Localization at heat-shock puffs

We used the antibodies to HSF to analyse its distribution in salivary gland polytene chromosomes, which provide a unique opportunity to visualize transcription factor-DNA interactions *in vivo*. Immunostaining^{28–30} of unspread polytene nuclei showed a diffuse distribution of HSF all over the chromosomes in the absence of heat shock, with a significant lack of staining of the surrounding cytoplasmic material (Fig. 3b). On a heat-shock stimulus at 36.5 °C for 25 min, this general diffuse pattern of staining was replaced by the appearance of staining at discrete sites scattered over the entire set of chromosomes (Fig. 3d).

The number and map positions of polytene loci that stain for HSF were determined using polytene chromosome spreads. Nine loci in *D. melanogaster*: 33B, 63BC, 64F, 67B, 70A, 87A, 87C, 93D and 95D have prominent puffing on heat shock³¹. Eight of these puffs are evident in Fig. 4a (arrows) as decondensed sections of the chromosome that show decreased staining for the DNA fluorescent dye bisbenzimidazole. In this chromosome spread, the heat-shock puff at 70A was obstructed by an overlapping chromosomal arm, and the two homologues on the distal right arm of chromosome 2 were asynapsed, as was a region of chromosome 3 between sections 86 and 98, which includes the four heat-shock puffs at 87A, 87C, 93D and 95D.

There was intense staining for HSF at the two major heat-shock puffs at 87A and 87C (Fig. 4b). These loci carry two and three copies, respectively, of the *hsp70* gene, each of which has two perfect and two imperfect HSEs. The highest degree of staining was consistently at 87C, which carries an additional eight copies of the heat-shock-inducible $\alpha\beta$ RNA repeat^{32,33}. Intense staining was also observed at the heat-shock locus 67B, site of the *hsp27*, *hsp26*, *hsp23* and *hsp22* genes. There were high amounts of HSF staining at 33B, 63BC (*hsp83*), 64F and 87A

(*hsp70*), and a low amount of staining at 70A and 95D (*hsp68*). The HSF staining intensity at each heat-shock locus seems to be roughly consistent with the estimated number of HSF-binding sites, or HSEs. It is notable that the staining for HSF at each puff site was not spread over the entire decondensed region, but restricted to a relatively sharp band, indicating that the register between the thousand or so individual chromatids is maintained across the puffed region. The superposition of HSF staining at puffs is evident from a double exposure for HSF and DNA (data not shown).

In comparison to the staining of the eight heat-shock puffs discussed above, the prominent heat-shock puff at 93D consistently stained poorly for HSF, although the staining was above background (see below). The 93D locus is known to possess unusual properties that set it apart from the other heat-shock loci, the most relevant difference being its unique induction by a variety of reagents that do not induce puffing at other heat-shock loci^{34,35}. Sequences upstream of the gene encoded at 93D (*hsw*) contain only one poor match to the HSE consensus sequence³⁶. In this context, the weak staining at 93D suggests

TABLE 1 HSF-binding sites on polytene chromosomes

1(X)	2L	2R	3L	3R
1 C ++	21 B +	42 C +	61 C +	84 E ++
2 A $+\frac{1}{2}$	B $+\frac{1}{2}$	43 C $+\frac{1}{2}$	F $+\frac{1}{2}$	E ++
F +	23 A +	E ++	62 B $+\frac{1}{2}$	F ++
3 AB +	C +	44 E +	63 BC +++	F ++
5 C ++	F +	F $+\frac{1}{2}$	64 F +++	85 B $+\frac{1}{2}$
E $+\frac{1}{2}$	24 B +	F $+\frac{1}{2}$	65 A $+\frac{1}{2}$	C +
E $+\frac{1}{2}$	D $+\frac{1}{2}$	45 D +++	DE $+\frac{1}{2}$	D +
6 B +	25 C +	46 B +	66 B ++	E $+\frac{1}{2}$
7 D +	D +	47 A $\frac{3}{4}$	B ++	86 C +
D +	26 AB ++	B $\frac{3}{4}$	C ++	D +
E +	D +	C $\frac{3}{4}$	C ++	F +
8 C +	28 A +	D $\frac{3}{4}$	CD ++	87 A +++
F $\frac{3}{4}$	D +	48 E $+\frac{1}{2}$	CD ++	C +++++
9 B $+\frac{1}{2}$	F +	49 D +	D $+\frac{1}{2}$	E ++
10 A +	29 E +	E +	67 B +++++	88 D +
A +	30 D +	50 B +	C ++	EF $+\frac{1}{2}$
B ++	33 B +++	D +	C ++	89 D +
E $+\frac{1}{2}$	34 A +	E +	E +	91 CD +
12 B +	AB ++	51 A +	E +	93 C $+\frac{1}{2}$
E +	C +	53 B +	68 A +	94 B +
F +	36 A $+\frac{1}{2}$	54 B ++	70 A +	C +
13 A +	F +	E +	C +	95 C +
E ++	37 B +	F $+\frac{1}{2}$	71 C +	D $+\frac{1}{2}$
F +	D +	55 B $+\frac{1}{2}$	72 C +	F +
14 D +	F $+\frac{1}{2}$	C +	73 A +	F +
15 DE $+\frac{1}{2}$	38 A +	56 BC +	C +	96 A +
19 A +	B +	D ++	C +	97 D +
	C +	57 C +	E $+\frac{1}{2}$	99 E +
	D +	F $+\frac{1}{2}$	E $+\frac{1}{2}$	F +
	F +	F $+\frac{1}{2}$	74 E ++	F $+\frac{1}{2}$
	39 A +	58 E +	75 AB +++	100 B +
	D +	59 E $\frac{3}{4}$	F +	B +
	40 E $+\frac{1}{2}$	F $+\frac{1}{2}$	76 E $+\frac{1}{2}$	D ++
	E $+\frac{1}{2}$	60 E +	77 B +	
	F $+\frac{1}{2}$	E +		

164 positive sites were scored by visual estimation of the fluorescence intensity, and ranked on a scale of $\frac{3}{4}$ to +++++. The appearance of a $\frac{1}{2}$ after a plus sign signifies an intensity between two values (for example, $+\frac{1}{2}$ is between + and ++). An additional 39 sites not tabulated showed weak but reproducible staining above background and were scored as $\frac{1}{2}$; the heat-shock locus 93D is included in this group. The fourth chromosome was not analysed. Map positions and degree of staining were independently confirmed using two or more chromosome spreads. The locations of the stained bands were extremely consistent between chromosome spreads, and the intensity of each fluorescent band was reproducible, although variation within a small range (one +) occurred at some loci. The presented values represent a mean of at least two readings. The relationship between an assigned value and the fluorescent intensity can be gauged by referring to the marked sites in Fig. 4d between chromosome sections 61 and 67. The locations of major heat-shock genes (in bold type) are 63BC (*hsp83*), 67B (*hsp23*, 26, and 27), 87A and 87C (*hsp70*), 95D (*hsp68*), 33B, 64F and 70A.

that other factors in addition to HSF may be involved in the regulation of this locus.

HSF binding to multiple loci

Besides staining at the prominent heat-shock puffs, HSF staining was detected at an unexpectedly large number (156) of polytene sites located throughout the euchromatic chromatin (Table 1). The intensity of staining at the 156 sites generally covered a broad range, with 31 sites being stained at intermediate levels (++) or (+++), similar to the staining of several heat-shock puffs. An additional 39 sites (including the 93D heat-shock puff) showed weak, but consistent staining above background. These sites were not included in Table 1. An example of all sites tabulated between chromosome sections 61 and 67 on the left arm of the third chromosome is shown in Fig. 4c, d.

What could be the relevance of HSF binding to so many loci? Some of the loci could represent minor heat-shock-inducible genes. Indeed, 48E and 88EF are minor heat-shock puff sites whose appearance is somewhat variable and dependent on developmental state³¹. 39D, 54E, 61F, 62B, 85B, 85D, 86C, 88E and 100D sites correspond to sites that show some hybridization to cDNAs or messenger RNAs from heat-shocked cells^{37,38}. The 39D and 88E sites also clearly accumulate RNA polymerase II after heat shock³⁹. Among the aforementioned loci, 39D includes the histone genes, whose transcription is maintained during heat shock⁴⁰, and 88E includes the heat-shock cognate gene *hsc4*, whose steady state RNA level is induced twofold after heat shock⁴¹. Intriguingly, five out of the six constitutively expressed heat-shock cognate genes⁴² map to chromosomal sites that have specific HSF binding: *hsc1* (70C), *hsc3* (10E), *hsc4* (88EF), *hsc5* (50E) and *hsc6* (5C), but not *hsc2* (87D). The *hsc3* and *hsc4* genes have been sequenced; there are two 6/9 matches to the HSE consensus upstream of *hsc4* (ref. 43), and *hsc3* had an 8/9 match to the HSE consensus sequence at -66 (D. Rubin and K. Palter, personal communication). Overall, these findings suggest a role for HSF-mediated transcription of the heat-shock cognates as an integral part of the chromosomal response to heat stress. No HSF binding was observed at 55A, the cytological location of the *HSF* gene itself, indicating the absence of transcriptional autoregulation.

In addition to the minor heat-shock loci, some of the chromosomal sites staining for HSF probably reflect the accessibility of individual HSEs or HSE-like sequences that occur statistically in the *Drosophila* genome. These nonfunctional sites could include the 39 weakest sites not tabulated as well as the weaker of the 156 tabulated sites. But those sites that stain by amounts approaching the traditional heat-shock loci should carry multiple elements for HSF binding, and it is these elements that are likely to be of functional significance.

Interaction with developmental loci

Two prominent HSF-binding sites at 74E and 75AB are coincident with the major ecdysone-inducible puff sites at 74EF and 75B. The 74EF/75B puffs are among six early developmental puffs induced by the insect steroid hormone, ecdysone, at the onset of metamorphosis⁴⁴. These puffs can be seen on the unshocked, late third instar larval chromosome; they show no preferential accumulation of HSF (Fig. 4e, f). On heat shock, both puffs, like all other puffs normal to development, undergo regression. HSF relocalized to two highly discrete sites that map to 74E and 75AB (Fig. 4h). In some chromosomal preparations, diffuse staining was also observed in the decondensed region between 74E and 75AB, indicating HSF binding at the puffs before complete regression (compare Fig. 4b, d, h). The localization of HSF to 74E/75AB is not an artefact of indiscriminate HSF interactions with chromosome puff material; no staining, for example, was detectable at a neighbouring developmental puff at 78D (Fig. 4e-h). In addition to 74EF/75B, 47 out of 125 loci documented by Ashburner⁴⁴ to puff during the last day of larval life share map positions with HSF-binding sites.

A role for HSF-mediated repression?

The repression of normal cellular transcription during heat stress has been known since the earliest studies of the heat-shock response, but essentially all transcriptional studies on the system have been focused on the activation of heat-shock genes. It has been a tacit assumption that the repression of normal transcription is due to some general, stress-induced defect of the transcriptional machinery. Without excluding this possibility, our results suggest an alternative mechanism for repression, on the basis of direct interactions with HSF at key control loci. The 74EF/75B 'early' puffs encode *ets*-related^{45,46} and steroid receptor⁴⁷ DNA-binding proteins required for the sequential activation of about 100 'late' puffs, whose products are thought to act as realizator proteins that define the morphological properties of hormone-responsive tissues⁴⁸. By directly binding to a finite number of key regulator loci like 74EF/75B, HSF could effectively shut down entire developmental programmes until environmental conditions became more conducive to proper growth and development.

Discussion

We have described physical and cytological properties of the normal and induced forms of *Drosophila* HSF showing that the inactive form of HSF exists in the cell nucleus as a small oligomer which is converted to an active multimer on heat shock. As contacts with all three (nGAAn) modules of the HSE are required for high-affinity binding²², a decrease in the oligomeric state would suffice to decrease the number of contacts, and hence the affinity for the HSE. How might the oligomeric switch between the two forms of HSF be mediated? In addition to the three conserved leucine zipper motifs found in the N-terminal regions of HSF, a fourth zipper motif is uniquely found in the C-terminal region of *Drosophila* and mammalian HSF sequences^{12,15-17}. As these proteins, unlike the *S. cerevisiae* and *K. lactis* HSFs, are regulated for DNA binding, it is conceivable that the C-terminal leucine zipper motif could be involved in the formation of the inactive HSF oligomer by intermolecular and intramolecular interactions, or by interactions with a protein inhibitor. The relative thermostability of alternate zipper interactions would thus form the basis of the oligomeric switch.

The visualization of HSF bound to the heat-shock puff sites links earlier discoveries¹ with more recent molecular studies of the heat-shock regulator in *Drosophila*^{12,25,49-53}. The immunostaining experiments also revealed two unforeseen findings. First, the localization of the inactive HSF oligomer in nuclei of cultured cells and on polytene chromatin indicates that HSF is prelocalized in the nucleus to minimize the response time to the stress signal. How this nuclear and chromatin prelocalization is established and maintained is unknown. Second, the number of HSF-binding sites *in vivo* exceeds the known heat-shock genes, both major and minor. The identification of overlap between HSF-binding sites and key developmental loci such as the ecdysone-inducible *ets*-related and steroid receptor genes at 74EF/75B suggests a new function for HSF in the heat-shock-induced repression of normal gene activity.

How could HSF binding lead to activation of some genes and to repression of others? Potential differences in the HSF-binding sequence between positive and negative *cis*-elements could result in structural alterations that change the activity of the transcription factor. HSF could bind directly to other positive regulatory proteins and alter their activities. Or HSF could cause repression by steric hindrance of other transcriptional regulators or RNA polymerase II, if a high-affinity binding site were to overlap with an enhancer element, or were located in the path of the transcribing enzyme. The availability of an extensive set of clones for the 74EF/75B loci⁴⁵⁻⁴⁷ provide a unique opportunity to test these models and to determine the functional importance of HSF localization at nonheat-shock loci. □

Received 30 July; accepted 17 September 1991.

1. Morimoto, R. I., Tissieres, A. & Georgopoulos, C. (eds) *Stress Proteins in Biology and Medicine* (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1990).
2. Ritossa, F. *Experientia* **18**, 571–573 (1962).
3. Pelham, H. R. B. in *Stress Proteins in Biology and Medicine* (eds Morimoto, R. I., Tissieres, A. & Georgopoulos, C.) 287–300 (Cold Spring Harbor Laboratory, Cold Spring Harbor, New York, 1990).
4. Ellis, R. J. & van der Vies, S. M. *A. Rev. Biochem.* **60**, 321–347 (1991).
5. Pelham, H. R. B. *Cell* **30**, 517–528 (1982).
6. Amin, J., Ananthan, J. & Voellmy, R. *Molec. cell. Biol.* **8**, 3761–3769 (1988).
7. Xiao, H. & Lis, J. T. *Science* **239**, 1139–1142 (1988).
8. Wu, C. in *Stress Proteins in Biology and Medicine* (eds Morimoto, R. I., Tissieres, A. & Georgopoulos, C.) 429–442 (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York, 1990).
9. Sorger, P. *Cell* **65**, 363–366 (1991).
10. Sorger, P. K. & Pelham, H. R. B. *Cell* **54**, 855–864 (1988).
11. Wiederrecht, G., Seto, D. & Parker, C. S. *Cell* **54**, 841–843 (1988).
12. Clos, J. *et al. Cell* **63**, 1085–1097 (1990).
13. Scharf, K.-D., Rose, S., Zott, W., Schoff, F. & Nover, L. *EMBO J.* **9**, 4495–4501 (1990).
14. Jakobsen, B. K. & Pelham, H. R. B. *EMBO J.* **10**, 369–375 (1991).
15. Rabindran, S. R., Giorgi, G., Clos, J. & Wu, C. *Proc. natn. Acad. Sci. U.S.A.* **88**, 6906–6910 (1991).
16. Schuetz, T. J., Gallo, G., Sheldon, L., Tempst, P. & Kingston, R. E. *Proc. natn. Acad. Sci. U.S.A.* **88**, 7911–7915 (1991).
17. Sarge, K. D., Zimarino, V., Holm, K., Wu, C. & Morimoto, R. I. *Genes Dev.* (in the press).
18. Larson, J. S., Schuetz, T. J. & Kingston, R. E. *Nature* **335**, 372–375 (1988).
19. Mosser, D. D., Kottbauer, P. T., Sarge, K. D. & Morimoto, R. *Proc. natn. Acad. Sci. U.S.A.* **87**, 3748–3752 (1990).
20. Becker, J., Mezger, V., Courgeon, A.-M. & Best-Belpomme, M. *Eur. J. Biochem.* **189**, 553–558 (1990).
21. Zimarino, V., Wilson, S. & Wu, C. *Science* **249**, 546–549 (1990).
22. Perisic, O., Xiao, H. & Lis, J. T. *Cell* **59**, 797–806 (1989).
23. Sorger, P. K. & Nelson, H. C. M. *Cell* **59**, 807–813 (1989).
24. Andersson, L. O., Borg, H. & Mikaelsson, M. *FEBS Lett.* **20**, 199–201 (1972).
25. Wu, C. *et al. Science* **238**, 1247–1253 (1987).
26. King, W. J. & Greene, G. L. *Nature* **307**, 745–747 (1984).
27. Welshons, W. V., Lieberman, M. E. & Gorski, J. *Nature* **307**, 747–749 (1984).
28. Silver, L. M. & Elgin, S. C. R. *Proc. natn. Acad. Sci. U.S.A.* **73**, 423–427 (1976).
29. Jamrich, M., Greenleaf, A. L. & Bautz, E. K. F. *Proc. natn. Acad. Sci. U.S.A.* **74**, 2079–2083 (1977).
30. Zink, B. & Paro, R. *Nature* **337**, 468–471 (1989).
31. Ashburner, M. *Chromosoma* **31**, 356–376 (1970).
32. Lis, J., Ish-Horowitz, D. & Pinchin, S. M. *Nucleic Acids Res.* **9**, 5297–5310 (1981).
33. Mason, P. J., Torok, I., Kiss, I., Karch, F. & Udvary, A. *J. molec. Biol.* **156**, 21–35 (1982).
34. Lakhota, S. C. *J. Genet.* **66**, 139–157 (1987).
35. Bonner, J. J. & Pardue, M. L. *Cell* **8**, 43–50 (1976).
36. Hovemann, B., Walldorf, U. & Ryseck, R.-P. *Molec. Gen. Genet.* **204**, 334–340 (1986).
37. Lis, J. T., Neckameyer, W., Dubensky, R. & Costlow, N. *Gene* **15**, 67–80 (1981).
38. Spradling, A., Pardue, M. L. & Penman, S. *J. molec. Biol.* **109**, 559–587 (1977).
39. Bonner, J. J. & Kerby, R. L. *Chromosoma* **85**, 93–108 (1982).
40. Pardue, M. L., Kedes, L. H., Weinberg, E. & Birnstiel, M. L. *Chromosoma* **63**, 135–152 (1977).
41. Palter, K. B., Watanabe, M., Stinson, L., Mahowald, A. P. & Craig, E. A. *Molec. cell. Biol.* **6**, 1187–1203 (1986).
42. Lindquist, S. & Craig, E. A. *Rev. Genet.* **22**, 631–677 (1988).
43. Perkins, L. A. *et al. Molec. cell. Biol.* **10**, 3232–3238 (1990).
44. Ashburner, M. in *Results and Problems in Cell Differentiation* Vol. 4 (eds Beermann, W. *et al.*) 101–151 (Springer, New York, 1972).
45. Burtis, K. C., Thummel, C. S., Jones, C. W., Karim, F. D. & Hogness, D. S. *Cell* **61**, 85–99 (1990).
46. Urness, L. D. & Thummel, C. S. *Cell* **63**, 47–61 (1990).
47. Seagraves, W. A. & Hogness, D. S. *Genes Dev.* **4**, 204–219 (1990).
48. Thummel, C. S. *Bioessays* **12**, 561–568 (1990).
49. Wu, C. *Nature* **309**, 229–234 (1984).
50. Wu, C. *Nature* **311**, 81–84 (1984).
51. Wu, C. *Nature* **317**, 84–87 (1985).
52. Parker, C. S. & Topol, J. *Cell* **36**, 357–369 (1984).
53. Wiederrecht, G., Shuey, D. J., Kibbe, W. & Parker, C. S. *Cell* **48**, 507–515 (1987).
54. Dignam, A. D., Martin, P. L., Shastry, B. S. & Roeder, R. G. *Meth. Enzym.* **191**, 582–598 (1983).
55. Bridges, C. B. *J. Hered.* **26**, 60–64 (1935).
56. Lefevre, G. in *The Genetics and Biology of Drosophila* Vol. 1a (eds Ashburner, M. & Novitski, E.) 31–66 (Academic, New York, 1976).

ACKNOWLEDGEMENTS. We thank M. L. Pardue and J. Bonner for comments on chromosome mapping, K. Palter for unpublished *hsc* sequences, S. Lindquist for a cell-staining protocol, and V. Zimarino and members of our laboratory for comments. J.T.W. was supported by the Alberta Heritage Foundation for Medical Research.

LETTERS TO NATURE

Formation of a planet orbiting pulsar 1829–10 from the debris of a supernova explosion

D. N. C. Lin, S. E. Woosley & P. H. Bodenheimer

University of California Observatories/Lick Observatory,
Board of Studies in Astronomy and Astrophysics, University of California,
Santa Cruz, California 95064, USA

THE 10-Earth-mass planet¹ in a nearly circular 0.7-AU orbit around PSR1829–10 is unlikely to have survived the supernova, or especially the pre-supernova evolution of the star that became the pulsar. Here we describe how the planet might have been created inside the young supernova remnant^{1–3}. The principal difficulty lies not in providing enough mass or conducive thermodynamic conditions for planet formation, but in explaining the large angular momentum ($\sim 3 \times 10^{48}$ erg s) and small eccentricity (< 0.1) of the orbit. We propose that the planet formed from a rotationally supported disk of ~ 0.02 solar mass of heavy elements that fell back from the supernova explosion to an initial radius of about 1,000 km. Viscous evolution of the disk then concentrated most of its angular momentum into a small amount of material at the disk's outer extremity: 10 Earth masses at 10^{13} cm. Here, dust grains that had condensed and precipitated towards the mid-plane grew through cohesive collisions and gravitational instabilities into 100-km planetesimals, which coagulated into the planet on a million-year timescale. We find the presence of a second planet, more massive and more distant, unlikely, although residual planetesimals may provide the fuel for γ -ray bursts.

Survival of a pre-supernova planet would encounter two main problems. First, and most serious, an orbit of $a_p = 0.7$ AU would place the planet inside the photosphere of any red supergiant likely to become a Type II supernova. Such stars typically have radii of 2–4 AU (ref. 4) for at least several hundred thousand years. Even the blue supergiant progenitor of SN1987A, which exploded with a radius of only 0.2 AU, lived for an extended

period⁵ as a red supergiant before shrinking to its smaller pre-supernova state. Planets formed with the main-sequence star would be enveloped during the star's later stages of evolution and would spiral inwards to evaporation. Second, the loss during the explosion of more than half the total mass of the system would make it hard for the system to remain bound in a circular orbit¹.

We therefore prefer to consider a scheme that bears considerable similarity to current models for ordinary planet formation, for example in the Solar System⁶, but operates in the unusual environment of the interior of a young supernova remnant. We shall refer to the planet as 'Phoenix', as it arises from the ashes of a dead star.

The supernova explosion typically occurs in a star with original mass 15–20 $M_\odot$. Material inside $\sim 1.6 M_\odot$ collapses and becomes the neutron star. Most of the rest is ejected, but a small amount falls back. Estimates of the fallback mass range from ~ 0.001 to $0.1 M_\odot$, depending on details of the poorly known explosion mechanism and the evolutionary state of the star^{7,8}. We assume that a few hundredths of $M_\odot$ is a reasonable value. The fallback probably occurs within the first few hours to days of the explosion, when the 'reverse shock', formed as the core ran into the envelope, arrives back at the centre of the supernova. The important question is then whether the fallback material has sufficient angular momentum to account for that of the planet, $J \approx 3 \times 10^{48}$ erg s.

The progenitor star, a B0 star on the main sequence (17 $M_\odot$), has a typical rotation period of 2 days⁹. At the mass corresponding to that of the fallback material, (T. A. Weaver and S.E.W. manuscript in preparation) the specific angular momentum $j \approx 10^{17}$ cm² s⁻¹. But as the star evolves through its multiple stages of nuclear burning, convection transports angular momentum out of the central region. The evolutionary study of a rotating 10 $M_\odot$ star¹⁰, with angular momentum transport, leads to an estimate that the immediate presupernova star will have $j \approx 5 \times 10^{16}$ cm² s⁻¹ at the silicon shell (which corresponds to the fallback mass) and $j \approx 10^{17}$ cm² s⁻¹ at the oxygen shell somewhat farther out. Such a large angular momentum in the silicon shell and iron core would cause difficulties during core collapse and

could lead to a deformed neutron star, gravitational radiation and the formation of a millisecond pulsar. Although such results cannot be ruled out (for example where the planet's angular momentum is extracted from the pulsar; see below), we prefer to follow a more conventional path and assume that the various stages of convection were efficient at transporting angular momentum out of the inner core.

Note, however, that the mass that falls back is not just from the layer that was directly above the iron core when it collapsed to a neutron star. During the explosion and reverse shock there is mixing¹¹⁻¹³, with appreciable amounts of material falling back even from the hydrogen shell¹³. Thus an average $j = 2 \times 10^{17} \text{ cm}^2 \text{ s}^{-1}$ in the fallback material is not unreasonable, especially in what may be a very rare case. Then $J \approx 10^{49} \text{ erg s}$, enough to account for that of the planet. It is the combination of relatively large fallback mass and angular momentum that makes the formation of planets larger than the Earth possible. The coalescence of a binary white dwarf progenitor¹⁴ and the ablation of a companion¹⁵ can also provide similar conditions.

The fallback material, with mass $\approx 0.02 M_\odot$ and angular momentum $J \approx 10^{49} \text{ g cm}^2 \text{ s}^{-1}$, forms a keplerian disk at a radius $\sim 10^8 \text{ cm}$ where gravity and centrifugal force balance¹⁶. It is composed of a mixture of Fe, Si, O, He and perhaps even a trace of H. The disk then evolves viscously¹⁷, transferring angular momentum outwards and much of the matter inwards. Matter near the outer boundary of the disk (r_{max}) moves out on a diffusion timescale $\tau_\nu \approx r^2/\nu$, where ν is the viscosity¹⁸. The total angular momentum, $J = 2\pi \int \Sigma \Omega r^3 dr$, where Σ is the surface density and Ω is the angular velocity, is conserved. J is mostly likely to be stored near r_{max} so that the surface density needed to contain most of the disk angular momentum becomes $\Sigma_d \approx J/(\pi r_{\text{max}}^2 j(r_{\text{max}})) \approx J/(\pi \Omega(r_{\text{max}}) r_{\text{max}}^4) = 10^3 J_{49} m^{-1/2} r_{13}^{-5/2} \text{ g cm}^{-2}$, where $J_{49} = J/10^{49} \text{ g cm}^2 \text{ s}^{-1}$, $m = M/M_\odot$ (where M is the neutron star mass), and $r_{13} = r_{\text{max}}/10^{13} \text{ cm}$. The disk mass becomes $M_d \approx \pi \Sigma_d r_{\text{max}}^2 \approx 3 \times 10^{29} J_{49} m^{-1/2} r_{13}^{1/2} \text{ g}$. Thus, at a_p , there is sufficient residual gas in the disk to form planet Phoenix, provided $J_{49} \approx 1$.

The timescale, τ_ν , for the disk to diffuse from 10^8 to 10^{13} cm increases with r , because ν is unlikely to increase markedly with r , and therefore it can be evaluated at $r_{13} \approx 1$. The viscosity is determined by turbulence¹⁹, which could be induced by magnetic instabilities²⁰. During the final stage of evolution, when grains condense at a temperature $T_g \approx 1,000\text{--}2,000 \text{ K}$ near $r_{13} \approx 1$, the disk is intrinsically unstable against convection in the direction normal to the disk plane²¹. Using a self-consistent mixing length model or a linear normal mode analysis^{22,23}, we find $\nu \approx 10^{-2} c_s^2/\Omega$ where c_s is the speed of sound. Taking $\mu \approx 20$ and $T \approx T_g$, we find $c_s \approx 10^5 \text{ cm s}^{-1}$, $\nu \approx 10^{14} \text{ cm}^2 \text{ s}^{-1}$ and $\tau_\nu \approx 10^5 \text{ yr}$.

The pulsar age, given by the spin-down time $\tau_p = P/\dot{P} \approx 10^6 \text{ yr}$, where P is the period, suggests that there would be sufficient time for the disk to diffuse well beyond 10^{13} cm . But at $r_{13} = 1$, the gas is depleted as heavy elements condense efficiently into grains, and the turbulence stops. The disk temperature is determined by two processes: irradiation from the pulsar and viscous dissipation. From black-body theory, there is a critical radius, $r_1 = 2.3 \times 10^{12} (L/L_\odot)^{1/2} T_3^{-2} \text{ cm}$ (where $T_3 = T_g/10^3 \text{ K}$) outside which irradiation has little influence on the condensation of the grains. The present pulsar luminosity is $L \approx L_\odot$, but an empirical relation^{24,25} shows that $L \propto t^{-3/2}$ so that at a pulsar age of 10^5 yr , $r_1 \approx 10^{13} \text{ cm}$. In the other heating process, T is determined by an equilibrium between dissipation and radiative loss^{21,26}. With an assumed opacity of $\sim 10^3 \text{ cm}^2 \text{ g}^{-1}$ for the high metal abundance, T_3 is attained at a radius $r_2 \approx 3 \times 10^{13} m^{1/3} T_3^{-2} (\Sigma/10^3 \text{ g cm}^{-2})^{4/3} \text{ cm}$, outside which grains condense and viscous evolution is terminated. Taking $\Sigma(r_2) \approx \Sigma_d$, we find $r_2 = 1.3 \times 10^{13} J_{49}^{4/13} m^{-1/13} T_3^{-6/13} \text{ cm}$. Thus we do not expect the formation of a massive planet at $r \gg 10^{13} \text{ cm}$.

If the neutron star initially rotated considerably faster than its present rate, the magnetosphere-disk interaction would have transferred considerable angular momentum from the neutron

star to the disk. The interaction occurs near the Alfvén radius, r_a , where the magnetic stress is balanced by the viscous stress, implying $r_a \approx 0.5(\mu_m^4/GMM^2)^{1/7}$ where $\dot{M}$ is the mass transfer rate in the disk²⁷. The present magnetic moment of the neutron star (μ_m) is estimated²⁸ to be $\sim 10^{30} \text{ gauss cm}^3$. The disk provides an effective torque to spin down the neutron star²⁹ on a timescale $\tau_s \approx 10^4 \text{ yr}$ (ref. 30) until it corotates with the disk at r_a and becomes an accreting X-ray pulsar. During the final stages of disk evolution when $\Sigma \approx \Sigma_d$ and $r_{13} \approx 1$, the mass transfer rate is $\dot{M} \approx 3\pi \Sigma \nu \approx 10^{-8} M_\odot \text{ yr}^{-1}$, so that $r_a \approx 10^8 \text{ cm}$. At that time, grain precipitation leads not only to gas depletion, but also to the inward propagation of a condensation front such that $\dot{M}$ vanishes throughout the disk on a timescale $\sim 0.1 \tau_\nu \approx 10^4 \text{ yr} \ll \tau_s$. Then r_a continues to expand, but when $2\pi/P > \Omega(r_a)$, accretion onto the neutron star through the magnetosphere is terminated, and residual disk gas is driven outward³¹. When the gas density in the magnetosphere becomes sufficiently low^{32,33}, the neutron star evolves from an X-ray pulsar into a radio pulsar.

Grains at the outer edge of the disk undergo cohesive collisions, rapid ($\leq 10^3 \text{ yr}$) growth and sedimentation to the mid-plane^{34,35}. If their dispersion velocity is damped by the inelastic collisions to $< M_d \Omega r/M \approx 500 \text{ cm s}^{-1}$, the thin disk of grains becomes gravitationally unstable and forms planetesimals³⁶ with masses $m_p \approx 10^{-12} M_\odot$. Planetesimals then grow through collisional coagulation³⁷ on a very rapid timescale until most of M_d is in ~ 100 planetesimals with mass $\sim 10^{-6} M_\odot$. At that time their Roche radii ($r_R = (m_p/3M)^{1/3} r$) marginally overlap with each other during a synodic period³⁸.

Subsequent growth occurs in a runaway in which the most massive object coagulates with the low-mass field planetesimals faster than field planetesimals can merge³⁹. The low orbital eccentricity of Phoenix is a compelling evidence for runaway accretion. The runaway requires the Safronov number $\theta \equiv Gm_p/2\sigma^2 r_p > 1$ where m_p and r_p refer to the mass and radius, respectively, of the most massive planetesimal, and σ is the velocity dispersion of the low-mass field planetesimals. The field objects primarily interact with each other such that a dynamical equilibrium between elastic gravitational scattering and dissipational collision is established when $\sigma \approx v_e \propto m_i^{1/3}$, where v_e and m_i are their surface escape velocity and mass, respectively⁴⁰. Thus, $\theta \approx (m_p/m_i)^{2/3} > 1$. The collision timescale between planetesimals is⁴¹ $\tau_c = m_i/[16\pi^{1/2} \Sigma_d r_p^2 \Omega(1+\theta)]$. For $\theta > 1$, $\tau_c \propto m_p^{-4/3}$, so that the most massive planetesimal is also the most frequently colliding one. The growth timescale is $\tau_g = m_p/\dot{m}_p = m_p \tau_c/m_i \approx (\sigma^2/8\pi^{1/2} G \Sigma_d r_p) \Omega^{-1}$. Even if the field planetesimals are composed of a few Earth-size objects so that $\sigma \approx 10^6 \text{ cm s}^{-1}$, $\tau_g (\sim 10^5 \text{ yr}) < \tau_p (10^6 \text{ yr})$. For lower-mass field planetesimals, τ_g may be considerably less.

As the field planetesimals are depleted, the collision timescale increases to $\geq 10^6 \text{ yr}$. In the long term, the gravitational perturbation of Phoenix may lead to secular increases in their orbital eccentricity in a manner similar to that proposed for the comets⁴³. When their periastron distance is reduced below $r_3 = (9M/(4\pi\rho))^{1/3} \approx 10^{11} \text{ cm}$ (ρ is the density of the planetesimals), large planetesimals are tidally disrupted⁴³. The orbits of the residual small objects eventually become parabolic, and their collisions with the neutron star may lead to γ -ray bursts⁴⁴. With the luminosity inferred from the pulsar spin down rate, the Poynting-Robertson effect⁴⁵ induces the clearing of grains with $r_g \leq 1 \text{ mm}$ in $\leq 10^6 \text{ yr}$. Thus at present we would not expect there to be sufficient small grains to provide any significant scintillation of the short-wavelength pulsar signal. $\square$

Received 8 August; accepted 18 September 1991.

1. Ballos, M., Lyne, A. G. & Shemar, S. L. *Nature* **352**, 311–313 (1991).
2. Gaskell, C. M. *Nature* **338**, 121–122 (1989).
3. Nakamura, T. *KEK Preprint No.* 88–127, February (1989).
4. Weaver, T. A., Woosley, S. E. & Zimmerman, G. B. *Astrophys. J.* **225**, 1021–1029 (1978).
5. Woosley, S. E., Pinto, P. A. & Ensmann, L. M. *Astrophys. J.* **324**, 466–489 (1988).
6. Safronov, V. S. & Ruzmaikina, T. V. in *Protostars & Planets II* (eds Black, D. & Matthews, M.) 959–980 (University of Arizona Press, Tucson, 1985).

7. Chevalier, R. A. *Astrophys. J.* **346**, 847–859 (1990).
8. Woosley, S. E. *Astrophys. J.* **330**, 218–253 (1988).
9. McNally, D. *The Observatory* **85**, 166–169 (1965).
10. Endal, A. S. & Sofia, S. *Astrophys. J.* **220**, 279–290 (1978).
11. Herant, M. & Benz, W. 1991. *Harvard-Smithsonian Center for Astrophysics, Preprint No. 3169* (1991); *Astrophys. J.* (in the press).
12. Fryxell, B. A., Müller, E. & Arnett, W. D. *Astrophys. J.* **367**, 619–634 (1991).
13. Müller, E., Fryxell, B. & Arnett, W. D. *Max-Planck Institut für Physik und Astrophysik, Preprint MPA 586* (1991); *Astr. Astrophys.* (submitted).
14. Podsiadlowski, Ph., Pringle, J. E. & Rees, M. J. *Nature* **352**, 783–785 (1991).
15. Ruderman, M., Shaham, J. & Tavani, M. *Astrophys. J.* **336**, 507–518 (1989).
16. Terebey, S., Shu, F. H. & Cassen, P. *Astrophys. J.* **286**, 529–551 (1984).
17. Lüst, R. Z. *Naturforsch.* **A7**, 87–98 (1952).
18. Lynden-Bell, D. & Pringle, J. E. *Mon. Not. R. astr. Soc.* **168**, 608–637 (1974).
19. Pringle, J. E. *Ann. Rev. Astr. Astrophys.* **19**, 137–162 (1981).
20. Balbus, S. & Hawley, J. *Astrophys. J.* **376**, 214 (1991).
21. Lin, D. N. C. & Papaloizou, J. C. B. in *Protostars & Planets II* (eds Black, D. & Matthews, M.) 981–1072 (University of Arizona Press, Tucson 1985).
22. Lin, D. N. C. & Papaloizou, J. C. B. *Mon. Not. R. astr. Soc.* **191**, 37–48 (1980).
23. Røed, S. P., Papaloizou, J. C. B. & Lin, D. N. C. *Astrophys. J.* **329**, 739–763 (1988).
24. Itoh, N., Sato, N. & Adachi, T. *Preprint* (Sofia University, Tokyo, 1990).
25. Manchester, R. N. & Taylor, J. H. in *Pulsars* (Freeman, San Francisco, 1977).
26. Shakura, N. I. & Sunyaev, R. A. *Astr. Astrophys.* **24**, 337–355 (1973).
27. Treves, A., Maraschi, L. & Abramowicz, M. *Publ. Astr. Soc. Pacific* **100**, 427–451, (1988).
28. Gunn, J. E. & Ostriker, J. P. *Astrophys. J.* **160**, 979–1002 (1970).
29. Lamb, F. K., Pethick, C. J. & Pine, D. *Astrophys. J.* **184**, 271–289 (1973).
30. Ghosh, P. & Lamb, F. K. *Astrophys. J.* **234**, 296–316 (1979).
31. Illarionov, A. & Sunyaev, R. A. *Astr. Astrophys.* **39**, 185–195 (1975).
32. Goldreich, P. & Julian, W. H. *Astrophys. J.* **157**, 869–880 (1969).
33. Kennel, C. F., Fujimura, F. & Pellat, R. *Space Sci. Rev.* **24**, 407 (1979).
34. Weidenschilling, S. J. *Gerlands Beitr. Geophys.* **96**, 21–33 (1987).
35. Weidenschilling, S. J. & Cuzzi, J. N. in *Protostars & Planets III* (eds Levy, E. & Matthews, M.) (University of Arizona Press, Tucson, in the press).
36. Goldreich, P. & Ward, W. R. *Astrophys. J.* **183**, 1051–1061 (1973).
37. Safronov, V. S. *Evolution of the Protoplanetary Cloud and the Formation of the Earth and Planets* (Nauka, Moscow, 1969); (Engl. trans.) NASA TTF-677 (1972).
38. Wetherill, G. W. *Ann. Rev. Astr. Astrophys.* **18**, 77–133 (1980).
39. Wetherill, G. W. & Stewart, G. R. *Icarus* **77**, 330–357 (1989).
40. Stewart, G. R. & Wetherill, G. W. *Icarus* **74**, 542–553 (1988).
41. Binney, J. & Tremaine, S. *Galactic Dynamics* (Princeton University Press, 1987).
42. Harwit, M. & Salpeter, E. E. *Astrophys. J.* **186**, L37–39 (1973).
43. Sridhar, S. & Tremaine, S. *Icarus reprint* (Canadian Institute of Theoretical Astrophysics, 1991).
44. Tremaine, S. & Zytkov, A. *Astrophys. J.* **301**, 155–163 (1986).
45. Harwit, M. *Astrophysical Concepts* (Wiley, New York, 1973).

ACKNOWLEDGEMENTS. We thank J. Cuzzi, L. Hernquist and S. Ruden for discussions. This work was supported by NASA and the NSF. D.L. acknowledges support as a Guggenheim fellow.

Creation by stellar ablation of the low-mass companion to pulsar 1829–10

Julian H. Krolik

Department of Physics and Astronomy, Johns Hopkins University, Baltimore, Maryland 21218, USA

THE pulsar 1829–10 has a remarkable companion¹ of only ~10 Earth masses, which occupies a nearly circular orbit of 184-day period. Bailes *et al.*¹ speculate that this companion is a planet which has either survived the earlier stellar evolution and supernova which created the pulsar, or else formed through some sort of coagulation process during the lifetime of the pulsar. I argue here that another interpretation, which they excluded, may actually be more plausible: that the companion began its life as a star, and has been ablated down to its present mass by absorbing a portion of the pulsar's spindown energy. That similar phenomena have already been seen in two other binary pulsars, PSR 1957+20 (refs 2, 3) and PSR 1744–24A (refs 4, 5), lends preliminary credence to this suggestion. The (surprisingly small) final mass of the remnant is determined by the interplay between decreasing spin-down luminosity, recession of the companion from the pulsar as a result of its mass loss, and, most importantly, shrinkage of the companion due to convective cooling of its interior.

The fact that any companion remains bound after the supernova argues for the existence of a close binary. If there had been no companion, the remaining neutron star could only be more than half of the presupernova mass of the system if the mass immediately before explosion were $\approx 3M_{\odot}$. Because the initial mass of an isolated star that eventually makes a supernova is probably $>8M_{\odot}$ (ref. 6), the winds during its main-sequence and red-giant phases must somehow have carried away 2/3 of its initial mass, a rather daunting requirement. But there are several possible mechanisms involving close stellar companions that result in a bound system after the explosion⁷: mass transfer from the primary to the secondary during the primary's red-giant phase, spiral-in of a low-mass companion engulfed in the primary's red giant envelope, or accretion-induced collapse of a white dwarf.

Suppose, then, that this system began life as a close binary. Even if the orbit just before the explosion were circular, it would become eccentric as a result of the mass suddenly lost from the system. As Bailes *et al.* remark, explaining the low eccentricity of the present orbit is the hardest single problem for the planet interpretation. In the picture proposed here, however, one would expect a circular orbit. If the companion has a convective

envelope (as all main-sequence stars of less than $1.5M_{\odot}$ do), the characteristic time for the orbit to be made circular by tidal dissipation ($\approx e/(de/dt)$, where e is eccentricity) is less than the spin-down time (1.25×10^6 yr; ref. 1) for orbital separation D less than $\sim 2.4(M_c/M_{\odot})^{-7/32}R_c$, where R_c is the companion's radius and M_c is its mass⁸. Any additional dissipation would only hasten circularization, as circular orbits are the least-energy state for fixed angular momentum.

Ablation from the companion can also produce a circular orbit. If it takes place predominantly from the side of the companion facing the pulsar (as is likely), the companion's specific orbital energy increases at a rate $\sim |M/M_c|c_s^2$, where c_s is the sound speed at the sonic point of the ablation flow. At the same time, the companion's specific orbital angular momentum j can also increase, because the inner surface of the companion has smaller j than the mean whenever the companion rotates in the same sense as the orbit, or not too rapidly in the other direction:

$$\frac{1}{j} \frac{dj}{dt} = \frac{M_c}{M_c} [(1-s)(1-s\Omega_*/\Omega_{\text{orb}}) - 1] \quad (1)$$

where Ω_* is the star's rotation rate, $s = \phi_j(1 + M_c/M_n)R_c/D$ and M_n is the neutron star mass. Uncertainty about the exact distribution of ablation over the surface of the companion is included in the dimensionless parameter ϕ_j . When $\phi_j(R_c/D)(1 + \Omega_*/\Omega_{\text{orb}}) > (c_s/v_{\text{orb}})^2$, where v_{orb} is the orbital velocity, the orbital angular momentum increases faster than the energy, and the orbit becomes more nearly circular. This condition is likely to be met in a close binary, because R_c/D is not too small, and c_s falls considerably below the orbital escape speed once the companion loses a significant amount of mass. The material leaves the system when it is accelerated to the orbital escape speed by the ram pressure of the pulsar wind.

Controversy exists over the mechanism by which pulsar companions absorb a portion of the pulsar's spin-down power (see refs 9–13 for a selection of views). I avoid the issue of how such a low-frequency wave penetrates the plasma surrounding the companion, and simply adopt a scaling law:

$$\frac{dM_c}{dt} = -\frac{\phi_m}{4} \frac{R_c^3 L_{\text{sd}}(t)}{GM_c D^2} \quad (2)$$

where ϕ_m is a dimensionless free parameter.

The spin-down luminosity L_{sd} evolves as

$$L_{\text{sd}}(t) = \frac{L_0}{[1 + (2L_0/I\Omega_0^2)t]^2} \quad (3)$$

where L_0 is the initial luminosity, I is the neutron star's moment of inertia and Ω_0 is its initial rotation rate. Substantial mass loss

can be expected when the initial mass-loss timescale

$$t_{m0} = 4.78 \times 10^3 \phi_m^{-1} \left(\frac{M_c}{M_\odot} \right)^{-1} \left(\frac{D(0)}{R_\odot} \right)^2 \times \left(\frac{L_0}{10^{38} \text{ erg s}^{-1}} \right)^{-1} \text{ yr} \quad (4)$$

is shorter than or comparable to the initial spin-down timescale, that is, when

$$\frac{D(0)}{R_\odot} \lesssim 3.6 \phi_m^{1/2} \left(\frac{M_c(0)}{M_\odot} \right)^{1/2} \left(\frac{P_0}{10 \text{ ms}} \right)^{-1} \quad (5)$$

where P_0 is the initial pulsar spin period and a lower-main-sequence mass-radius relation $R \propto M$ has been assumed for the initial state of the companion.

At first, the cooling time of the star is much longer than t_{m0} so its immediate response is essentially adiabatic. When stars lose mass without losing heat, they expand with $R_c \propto M_c^{-1/3}$ and central temperature $T_c \propto M_c^{4/3}$. Nuclear reaction rates are so sensitive to temperature that a small mass loss effectively shuts them off. If the star continues to lose mass, its mean temperature eventually falls low enough for electrons and protons to recombine to neutral hydrogen. Again assuming that $R_c(0) \propto M_c(0)$, the mass at which recombination begins is $M_{\text{rec}} \approx 0.011 M_\odot$, independent of the initial mass of the star.

Unfortunately, to properly follow the evolution of the companion beyond this point requires detailed stellar evolutionary modelling which will not be considered here. The principal issue is the rate of cooling. Recombination occurs when the temperature is $\geq 10^4$ K, a regime in which the opacity is several orders of magnitude greater than Thomson, but falls steeply as the star continues to cool¹⁴. If radiative diffusion losses were the only cooling mechanism, the cooling time would still be much longer than the mass-loss time unless the opacity fell far below Thomson. But recombination also significantly decreases the adiabatic temperature gradient. The combination of considerable opacity and a shallow adiabatic temperature gradient could drive a large convective heat flux.

Sufficiently rapid cooling deflects the mass-radius relation from $R_c \propto M_c^{-1/3}$ down to the zero-temperature mass-radius relation for a hydrogen-helium mixture¹⁵ (see Fig. 1):

$$R \approx R_m \begin{cases} (M/M_m)^{0.2} & M \leq M_m \\ (M/M_m)^{-1/3} & M \geq M_m \end{cases} \quad (6)$$

where $R_m \approx 0.1 R_\odot$ at $M_m = 0.0026 M_\odot$. On the low-mass branch, support against gravity is primarily due to Coulomb repulsion of pressure-ionized atoms; on the high-mass branch, it is due to electron degeneracy pressure. Because the zero-temperature radius for a given mass is much smaller than the adiabatic radius and $\dot{M}_c \propto R_c^3$, rapid cooling can quickly quench mass loss.

While the companion loses mass, its orbit also changes. Following the argument that the companion's specific angular momentum increases more rapidly than its specific energy, suppose that the orbit stays circular. In that case, the orbital separation D changes according to:

$$\frac{dD}{dt} = -D \frac{\dot{M}_c}{M_c} \left(\frac{2\varepsilon + M_c/M_n}{1 + M_c/M_n} \right) \quad (7)$$

where $\varepsilon = 1 - (1-s)(1-s\Omega_*/\Omega_{\text{orb}})$. Two effects drive orbital expansion: the difference between the specific angular momentum of the ablated material and the mean for the companion; and the decreasing mass of the system. Because the former is $\propto s \propto R_c/D$, cooling has the secondary effect of also quenching orbital evolution.

In the discussion so far, no allowance has been made for any torques exerted on the departing gas by the binary's non-axisymmetric gravitational potential. These may be significant (indeed they may be the reason why the orbit of PSR1957+20 is shrinking rather than expanding¹⁶), but to evaluate their magnitude

requires a detailed hydrodynamic calculation, and I omit consideration of this effect for the time being.

The companion's spin—and hence ε —is determined by two competing effects. On the one hand, mass loss removes angular momentum and changes the companion's moment of inertia; on the other, tidal dissipation⁸ tends to keep the spin locked to the orbital frequency. Dissipation depends so sensitively on $D/R_c \propto (D/R_c)^{-6}$ that it sometimes dominates and is sometimes negligible.

It has already become clear that substantial mass loss only takes place when the initial orbit of the binary is rather close, no more than several stellar radii. If $M_c > M_n$, $D \propto M_c^{-1}$ (see equation (7)), so that for times short compared with the initial spin-down time, M_c decreases linearly with time. Smaller companion masses yield slower orbital evolution: $d \ln D / d \ln M_c \approx \varepsilon + M_c/M_n$ when $M_c < M_n$. This is the reason why the observed D is only $\sim 10^2 \times D(0)$ whereas the observed M_c is $\sim 10^{-4} M_c(0)$. Moreover, because $\dot{M}_c \propto M_c^{-2} D^{-2}$ during the adiabatic evolution phase, mass loss tends to be either minor (if the criterion of equation (5) is not satisfied) or catastrophic.

Three mechanisms have a role in stopping mass loss: falling spin-down luminosity, increasing orbital separation and, most important, stellar cooling. The last factor, as I remarked earlier, starts to act when $M_c \lesssim 0.01 M_\odot$. Defining the stellar cooling time by the time required to radiate away the recombination energy of its atoms (when recombination occurs, kT is rather smaller than 13.6 eV), cooling can cause the star to shrink fast enough to quench mass loss if

$$\left(\frac{D}{D_{\text{obs}}} \right)^2 \left(\frac{T_s}{10^4 \text{ K}} \right)^4 > 1.7 \times 10^{-2} \phi_m \left(\frac{M_c}{M_{\text{rec}}} \right)^{-1} \left(\frac{R_c}{R_\odot} \right) \times \left(\frac{L_{\text{sd}}}{10^{38} \text{ erg s}^{-1}} \right) \quad (8)$$

where T_s is the surface temperature, and D has been normalized to the observed separation.

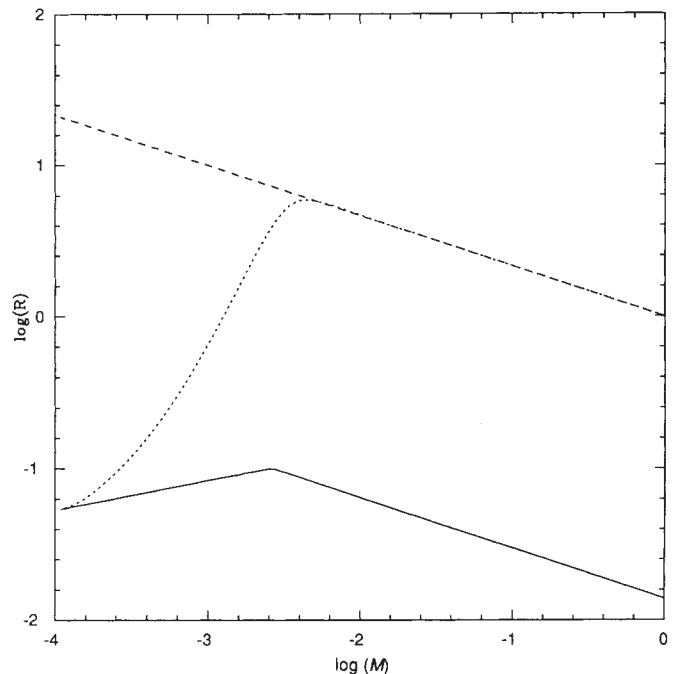


FIG. 1 Mass-radius relations in solar units for evaporating stars: the solid line shows an approximation to the zero-temperature mass-radius relation for hydrogen-helium mixtures found in ref. 15; the dashed line shows the adiabatic mass-radius relation for a star which began as a $1 M_\odot$ main-sequence star; the dotted curve shows a possible trajectory in this plane as a result of rapid convective cooling.

We may estimate T_s by assuming the entire star recombines at once, and making use of the fact that $d \ln T / d \ln p \approx 0.1$ in recombination zones¹⁷. In terms of σ_H , the mean opacity per proton at the photosphere, the temperature $T_s \approx 2.8 \times 10^4 (M_c / M_{\text{rec}})^{0.9} (R_c / R_\odot)^{-0.8} (\sigma_H / \sigma_T)^{-0.1} \text{K}$. In reality, T_s is probably smaller than this, but the estimate emphasizes how efficient convection can be.

When cooling is significant, D and L_{sd} change comparatively slowly (equations (3) and (7)). It is convenient to represent the mass-radius relation in that stage by $R_c \propto M_c^\alpha$, where α is roughly the ratio of the mass-loss timescale to the cooling timescale. For constant α , equation (2) has the approximate solution:

$$M_c \approx M_{\text{rec}} \left[1 + (3\alpha - 2) \frac{t}{t_{\text{mc}}} \right]^{-1/(3\alpha - 2)} \quad (9)$$

where $t_{\text{mc}} \approx 0.5 (L_{\text{sd}} / 10^{38} \text{ erg s}^{-1})^{-1} (M_c(0) / M_\odot)^{-2} (D / 10 R_\odot)^2 \text{ yr}$ is the mass-loss time at the point where rapid cooling begins. Catastrophic evaporation occurs when $\alpha < 2/3$; gradual mass loss which leaves a remnant occurs when $\alpha \geq 2/3$.

When the mass loss is gradual, the zero-temperature mass-radius relation is encountered when $R_c \approx R_m$, typically a shrinkage in radius of a factor of ~ 50 from the point at which cooling began (see Fig. 1). Because the mass-loss time is $\propto M_c^{2-3\alpha}$, gradual mass loss also implies that eventually the mass-loss time becomes longer than the spin-down time. Thus, the final mass of the companion can be estimated by its value at the point where the star's radius is near that given by the zero-temperature relation: $M_c \approx [R_m / R_c(0)]^{1/\alpha} [M_{\text{rec}} / M_c(0)]^{1/(3\alpha)} M_{\text{rec}}$. In the limit of the most rapid possible cooling ($\alpha \gg 1$), the final mass is $\approx 0.01 M_\odot$; at the break point between gradual and rapid mass loss ($\alpha = 2/3$), the final mass is $\sim 3 \times 10^{-5} (M_c(0) / M_\odot)^{-2} M_\odot$. Thus, cooling at the weaker end of the range that results in a

remnant produces a final mass within an order of magnitude of that observed, and stronger cooling could result in a more massive remnant.

These arguments demonstrate that the low-mass companion seen in this binary pulsar could have started out as a star, arriving in its present state as a result of ablation driven by the pulsar. Its physical state does resemble that of a (low-mass) jovian planet in the sense that its material is probably condensed pressure-ionized hydrogen and helium, and ionic Coulomb repulsion provides the main support against gravity¹⁵. But it arrived at that state by a path altogether different from the one we normally associate with planets.

Note added in proof: After submission of this paper, I received a preprint from F. Rasio, S. Shapiro and S. Teukolsky presenting a similar idea. $\square$

Received 9 August; accepted 19 September 1991.

1. Bailes, M., Lyne, A. G. & Shemar, S. L. *Nature* **352**, 311–313 (1991).
2. Fruchter, A. S., Stinebring, D. R. & Taylor, J. H. *Nature* **333**, 237–239 (1988).
3. Fruchter, A. S. *et al. Astrophys. J.* **351**, 642–650 (1990).
4. Lyne, A. G. *et al. Nature* **347**, 650–652 (1990).
5. Nice, D. J., Thorsett, S. E., Taylor, J. H. & Fruchter, A. S. *Astrophys. J.* **361**, L61–L63 (1990).
6. van den Bergh, S. *Astrophys. J.* **327**, 156–163 (1988).
7. Verbunt, F. in *Neutron Stars and their Birth Events* (ed. Kundt, W.) 179–218 (Kluwer, Dordrecht, 1990).
8. Zahn, J.-P. *Astr. Astrophys.* **57**, 383–394 (1977).
9. Ruderman, M., Shaham, J. & Tavani, M. *Astrophys. J.* **336**, 507–518 (1989).
10. Phinney, E. S., Evans, C. R., Blandford, R. D. & Kulkarni, S. R. *Nature* **333**, 832–834 (1988).
11. Michel, F. C. *Nature* **337**, 236–238 (1989).
12. Eichler, D. & Levinson, A. *Astrophys. J.* **335**, L67–70 (1988).
13. Krolik, J. H. & Sincell, M. W. *Astrophys. J.* **357**, 208–215 (1990).
14. Clayton, D. D. *Principles of Stellar Evolution and Nucleosynthesis*, **224** (McGraw-Hill, New York, 1968).
15. Zappol, H. S. & Salpeter, E. E. *Astrophys. J.* **158**, 809–813 (1969).
16. Ryba, M. & Taylor, J. H. (1991). *Princeton University preprint*.
17. Mihalas, D. *Stellar Atmospheres*, 187 (Freeman, New York, 1978).

ACKNOWLEDGEMENTS. This work was partially supported by NASA.

Holographic images of iodine atoms on a silver surface from electron emission patterns

Peijun Hu & D. A. King

Department of Chemistry, University of Cambridge, Cambridge CB2 1EW, UK

BARTON¹ has suggested that photoelectron interference patterns may be used directly as holograms to obtain atomic-resolution images of surface structures. Bulk structures have been obtained previously by this means from experimental patterns of high-energy Kikuchi (quasi-elastically scattered) and Auger electrons^{2,3}. Here we test the feasibility of this technique for determination of surface structures using Auger intensity patterns obtained^{4,5} from iodine chemisorbed on a pseudomorphous silver monolayer on Pt{111}. By direct numerical holographic inversion, we obtain three-dimensional images which show that iodine adatoms are located in hollows of 3-fold symmetry on the surface. The images yield the site symmetry with good atomic resolution in the surface plane, but suffer from poor resolution along the Ag–I axis. We anticipate that data with better angular resolution obtained at low temperatures would improve the spatial resolution of such images.

At present, most techniques in surface crystallography use a trial-and-error procedure to determine structures, low-energy electron diffraction being the most important of these techniques. Obviously, a direct route offered by atomic-resolution holography would have good potential. In 1948 Gabor⁶ described the concept of holography. He considered that this would make it possible to achieve resolution in electron micro-

scopy limited only by the wavelengths of the electrons, as subsequently discussed⁷. Szöke⁸ suggested that many electron diffraction patterns could be regarded as holograms. This has been demonstrated^{1,9} for photoelectron diffraction patterns and also claimed¹⁰ for diffuse LEED patterns. More recently, profiles of X-ray photoelectron spectra have been used^{11,12} to produce real-space images by calculation. Hu *et al.*¹³ have examined the effect of the interference of object waves in the holographic transform. So far, however, few experimental studies have appeared. Harp *et al.*^{2,3} measured patterns of high-energy Kikuchi and Auger electrons and were able to reconstruct three-dimensional images relating to an averaged bulk structure from these patterns, treating them as holograms. To test this idea experimentally for a surface structure, we chose published data^{4,5} from an epitaxial silver monolayer grown on Pt{111} with a chemisorbed iodine overlayer¹⁴. Using the spatial distribution of Auger electrons from the silver atoms (355 eV kinetic energy, primarily the $M_5N_{4,5}N_{4,5}$ and $M_4N_{4,5}N_{4,5}$ transitions) as a hologram, we show that real-space images can be obtained by direct numerical reconstruction which agree qualitatively with the results of previous structural analyses of a $(\sqrt{3} \times \sqrt{3}) R30^\circ$ iodine overlayer on bulk Ag{111}.

As an illustration of the holographic process for the I/Ag/Pt{111} 'sandwich' structure, an Auger electron emitted from a silver atom has two important paths to a detector: directly, without intermediate scattering from the overlying iodine layer (the reference wave); or after scattering by overlying iodine adatoms (the object waves). The resulting angular distribution intensity map is, in fact, the interference pattern of reference and object waves, which could in principle be regarded as a Fraunhofer hologram¹⁵. Two conditions for good Fraunhofer holography need to be considered, the first being that the initial state of the emission process from the atom is isotropic, and

the second being that the reference wave should be significantly larger in amplitude than the object waves¹³. The Auger emission intensity map from Ag on Pt{111} without chemisorbed I is essentially featureless¹⁶, indicating that the first of these conditions is met in the case of I/Ag/Pt{111}. In Auger emission from a selected species at 355 eV, the second condition is also likely to be met¹³.

Real-space images are calculated from the electron interference pattern by direct inversion, as originally suggested¹. We have not, however, removed the reference waves $I_0(k)$ as suggested by Barton¹. Experimentally measured intensities from an Auger intensity map are used here, whence the precise value of $I_0(k)$ is unknown. Failure to remove the reference wave will, however, only introduce additional intensity near the origin of the reconstructed images¹⁰. For the I/Ag/Pt{111} structure, the Auger intensity superposition arises entirely from a single layer of Ag atoms:

$$I_1 + I_2 + I_3 + \dots = \sum I_i \quad (1)$$

where the diffraction intensity resulting from Auger electrons from atom i in the silver layer is I_i . The reason we sum intensities rather than amplitudes is that Auger electron waves emitted from different atoms are incoherent. In the 'sandwich' structure, where we consider emission from the silver monolayer, most of the signal detected will arise from the idealized case in which directly emitted Auger electrons arrive at the detector and interfere with object waves scattered from the overlayer of chemisorbed iodine adatoms. Furthermore, as there is only a single atomic layer of iodine above the silver emitter, multiple scattering effects in the Auger electron waves propagating towards the detector through the iodine layer are expected to be minimal, and as forward focusing^{17,18} is dominant, the 'object' waves should be singly scattered primarily in the forward-focusing cone.

The map of Auger intensity distribution observed experimentally⁴ for I/Ag/Pt{111} is shown in Fig. 1a. Using a single scattering formalism, we have been able to simulate a pattern which closely resembles this pattern by placing I adatoms in three-fold hollow sites on Ag{111}; no agreement is obtained

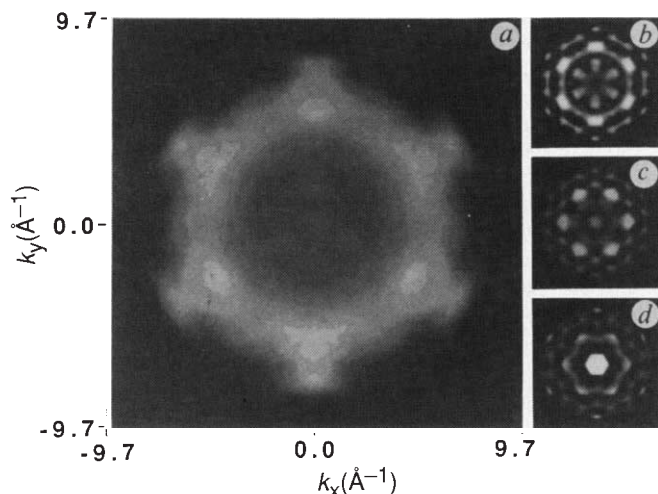


FIG. 1 *a*, The experimental map of Auger intensity distribution⁴ for I/Ag/Pt{111}, in which the Auger electrons are emitted from the Ag layer. The published data have been averaged using the 3-fold symmetry of the substrate, and projected into the k_x, k_y plane in reciprocal space, where k_x corresponds to the $\langle 10\bar{1} \rangle$ azimuthal direction and k_y to the $\langle \bar{1}21 \rangle$ direction. A linear grey-scale is used, in which white denotes large Auger electron intensity. *b*, Auger intensity map calculated assuming I adatoms in three-fold hollow sites. *c*, Map calculated with I in 2-fold bridged sites. *d*, Map calculated with I in on-top sites.

with I atoms in on-top or bridge sites. These results are also shown Fig. 1.

We display the real-space images obtained by holographic transform of the experimental data as two-dimensional intensity-distributions: a horizontal cut and a vertical cut through the three-dimensional image are shown in Figs 2 and 3. The origin of the real-space coordinate system is defined as being at the position of the source Ag atom, the x axis corresponds to the $\langle 10\bar{1} \rangle$ direction in the crystal, and the y axis corresponds to $\langle \bar{1}21 \rangle$ direction. We can see clearly that there are six bright patches in Fig. 2, which is a cut through a horizontal plane 2.4 Å above the Ag plane, where we anticipate finding the I atoms. As indicated in the figure, the bright patches occur along the $\langle \bar{1}21 \rangle$ azimuthal directions, strongly suggesting that I adatoms occupy 3-fold hollow sites. The image is inconsistent with on-top or bridge sites. Furthermore, the observation of 6-fold, and not 3-fold, symmetry is consistent with the occupation of both face-centred cubic (f.c.c.) and hexagonal close-packed (h.c.p.) sites on the surface. The observed real-space image (Fig. 2) is clearly a sum of all the atom positions shown in Fig. 4. Apart

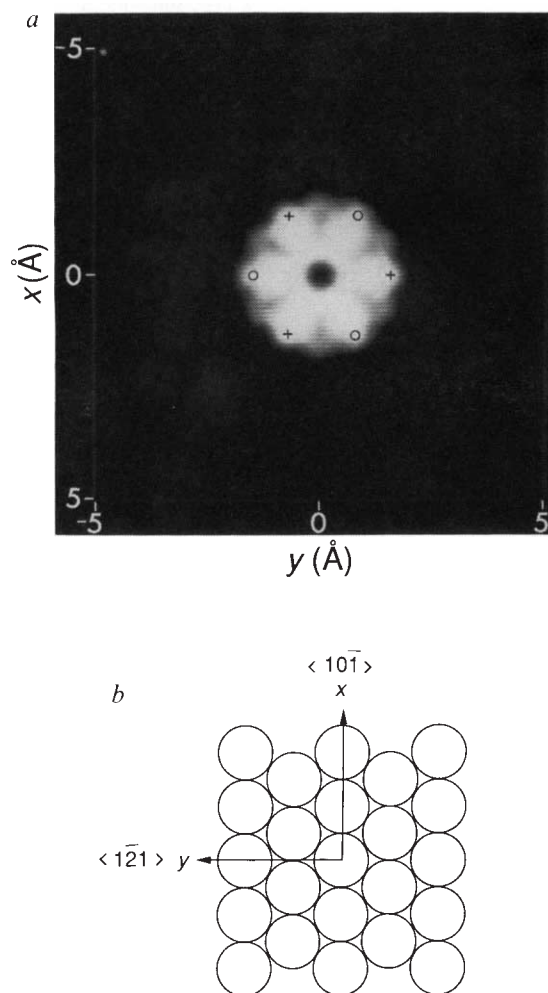


FIG. 2 *a*, A two-dimensional cut through the holographic image reconstructed from the experimental Auger intensity map shown in Fig. 1. The cut is taken along a plane parallel to the crystal surface and 2.4 Å above the emitter atoms. (+), Expected locations of I atoms in f.c.c. 3-fold hollow sites; (○) expected locations of I atoms in h.c.p. sites. The observed 6-fold symmetry arises because the pattern represents a simple summation (equation (1)) of all structures present on the surface, as observed from the position of each Ag atom in the surface. *b*, The crystallographic orientation of the Ag{111} pseudomorphic layer is demonstrated.

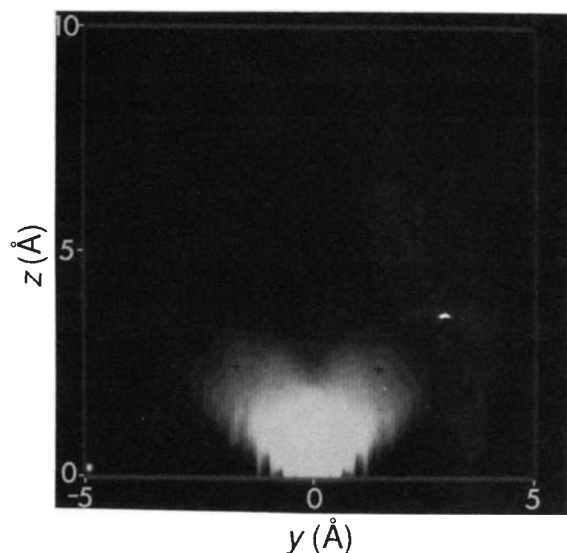


FIG. 3 A two-dimensional cut through the holographic image in the y - z plane orthogonal to the surface. (+), Expected locations of I atoms in three-fold hollow sites.

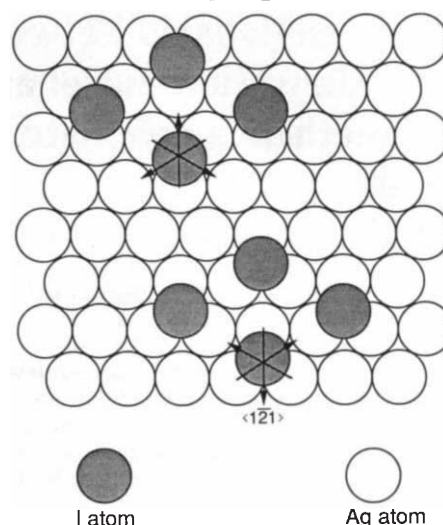


FIG. 4 Two kinds of iodine atoms on the Ag{111} surface corresponding to occupation of f.c.c. and h.c.p. sites. If only one were present the holographic image in Fig. 3 would show 3-fold symmetry; with both present the image is six-fold symmetric, as observed.

from the locations corresponding to the first-neighbour I atoms, no other features are imaged in Fig. 2. The intensity map is dominated by the interference fringes between the wave emitted from a silver atom and its waves forward-scattered from neighbouring iodine atoms. Figure 3 shows a vertical cut through the y - z plane along the $\langle 1\bar{2}1 \rangle$ azimuth. Compared with the horizontal cut in Fig. 2, the image in Fig. 3 is diffuse. This may arise in part from poor resolution in the z direction, a problem that is also common in optical holography. Nevertheless, it is clear that there are only two main beam-like patches, which correspond to an I-Ag bonding angle of $\theta \approx 33$ - 34° with respect to the surface normal. With I atoms in 3-fold hollow sites, this bond angle yields a Ag-I bond length of ~ 2.86 Å, in close agreement with analyses of I on Ag{111} by low-energy electron diffraction, photoelectron diffraction and SEXAFS (surface extended X-ray absorption fine structure)¹⁹⁻²¹. We note that these structural analyses also demonstrate that both f.c.c. and h.c.p. 3-fold sites are occupied, in agreement with the current work. But although the I/Ag/Pt{111} sandwich structure is substantiated by Auger data¹⁴, it should be noted that the low-energy electron diffraction pattern with one monolayer of Ag is not identical to that of the system I/Ag{111}¹⁴. The fractional order beams of the $(\sqrt{3} \times \sqrt{3})$ R30° structure show splitting, as observed for a $(\sqrt{3} \times \sqrt{3})$ R30° bromine overlayer on Ag{111}, where it is attributed to incommensuracy²². As I on bulk Ag{111}

is not incommensurate, and the structural evidence of the present work strongly supports occupancy of the high-symmetry 3-fold sites, it is likely that the incommensuracy is at the Ag/Pt{111} interface.

Frank *et al.*^{4,5} have provided a very different interpretation of their own data from that given here. From a shadowing model based on linear trajectories of electrons blocked by atoms in their path, they conclude that the I adatoms are in on-top sites. Our results of the holographic transform from the same experimental data are consistent with single scattering cluster calculations based on electron-scattering theory, illustrated in Fig. 1; both of these results strongly contradict their interpretation. Furthermore, crystallographic data do not support the notion that halogen adatoms would occupy on-top sites at low coverage on a {111} surface; the 3-fold hollow sites are always preferred^{23,24}.

We conclude that the real-space images reconstructed by direct numerical calculation from Auger intensity maps for the system I/Ag/Pt{111} are fully three-dimensional, showing the I adatoms located at threefold hollow sites. Structural information is obtained directly without the introduction of any parameters or variables. Resolution along the I-Ag axis is, however, poor. In the future, low-temperature, high-resolution data may be obtained, and the treatment improved, to provide better resolution. □

Received 20 June; accepted 17 September 1991.

1. Barton, J. *Phys. Rev. Lett.* **61**, 1356-1359 (1988).
2. Harp, G. R., Saldin, D. K. & Tonner, B. P. *Phys. Rev. Lett.* **65**, 1012-1015 (1990).
3. Harp, G. R., Daldin, D. K. & Tonner, B. P. *Phys. Rev. B* **42**, 9199-9202 (1990).
4. Frank, D. G., Batina, N., Golden, T., Lu, F. & Hubbard, A. T. *Science* **247**, 182-188 (1990).
5. Frank, D. G., Batina, N., McCarpair, J. W. & Hubbard, A. *Langmuir* **5**, 1141-1146 (1989).
6. Gabor, D. *Nature* **161**, 777-778 (1948).
7. Lichte, H. *Ultramicroscopy* **20**, 293-304 (1986).
8. Szöke, A. in *Short Wavelength Coherent Radiation: Generation and Applications* (eds Attwood, D. T. & Boker, J.) *AIP Conf. Proc. No. 147* (American Institute of Physics, New York, 1986).
9. Barton, J. *J. Electron Spectroscopy and Related Phenomena* **51**, 37-53 (1990).
10. Saldin, D. K. & de Andres, P. L. *Phys. Rev. Lett.* **64**, 1270-1273 (1990).
11. Wei, C. M., Zhao, T. C. & Tong, S. Y. *Phys. Rev. Lett.* **65**, 2278-2281 (1990).
12. Tong, S. Y., Wei, C. M., Zhao, T. C., Huang, H. & Li, H. *Phys. Rev. Lett.* **66**, 60-63 (1991).
13. Hu, P., Barnes, C. J. & King, D. A. *Chem. Phys. Lett.* **183**, 521-528 (1991).

14. Stickney, J. L., Rosasco, S. D., Song, D., Soriaga, M. P. & Hubbard, A. T. *Surf. Sci.* **130**, 326-347 (1983).
15. Born, M. & Wolf, E. *Principles of Optics* (Pergamon, Oxford, 1980).
16. Frank, D. G., Golden, T., Lu, F. & Hubbard, A. T. *MRS Bull.* **15**, 19 (1990).
17. Poon, H. C. & Tong, S. Y. *Phys. Rev. B* **30**, 6211-6213 (1984).
18. Orders, P. J. & Fadley, C. S. *Phys. Rev. B* **27**, 781-798 (1983).
19. Maglietta, M. *et al. Surface Sci.* **123**, 141-151 (1982).
20. Kang, W. M., Li, C. H. & Tong, S. Y. *Solid State Commun.* **36**, 149-154 (1980).
21. Citrin, P. H., Eisenberger, P. & Hewitt, R. C. *Phys. Rev. Lett.* **41**, 309-317 (1978).
22. Holmes, D. J., Panagiotides, N. & King, D. A. *Surf. Sci.* **222**, 285-295 (1989).
23. Farrell, H. H. in *The Chemical Physics of Solid Surface and Heterogeneous Catalysis* Vol. 3 (eds King, D. A. & Woodruff, D. P.) 225-264 (1984).
24. Lamble, G. & King, D. A. *Phil. Trans. R. Soc. Lond.* **A318**, 2032-217 (1985).

ACKNOWLEDGEMENTS. We thank A. T. Hubbard and D. G. Frank for digitized Auger maps from previously published work. The Oppenheimer Trust (UK) is acknowledged for a studentship to P.H.

Seasonal relationship between cloud condensation nuclei and aerosol methanesulphonate in marine air

G. P. Ayers & J. L. Gras

Division of Atmospheric Research, CSIRO, Private Bag 1, Mordialloc 3195, Australia

CHARLSON *et al.*¹ have suggested that cloud-droplet concentrations in remote marine regions might be indirectly controlled by dimethylsulphide (DMS) emissions from marine phytoplankton. Attempts to test the hypothesis that variations in DMS emissions lead to corresponding variations in marine aerosol composition and hence concentrations of cloud condensation nuclei (CCN) have so far proved inconclusive, primarily because of the inherent variability of the atmospheric species involved over the timescales typical of individual field measurements². Here we present nine years of data from Cape Grim (41° S) which show that there is a significant seasonal relationship between cloud condensation nuclei and methanesulphonate, an easily sampled oxidation product of DMS, but the relationship is nonlinear. These results confirm that DMS emissions strongly influence CCN concentrations, but we note that at low concentrations of methanesulphonate, there are indications that there may be another source of CCN, apart from DMS.

Individual field measurements have typically had timescales of hours. Recently, however, we have been able to establish a clear link between DMS and the aerosol components methanesulphonate (MSA) and non-sea-salt sulphate (n.s.s.-SO₄²⁻) in marine air at Cape Grim by taking longer averaging times³. We continue that approach here, seeking to identify any relationship between CCN and the DMS oxidation product MSA at Cape Grim as a function of the large seasonal variation known to occur in MSA (and DMS³) concentration at this site. We are constrained to this indirect comparison, rather than direct comparison of CCN and DMS concentrations, by the lack of a long-term DMS record paralleling the CCN record, which dates⁴ from 1981.

Data for CCN⁴ and high-volume aerosol MSA^{5,6} extending from 1981 to 1989 are shown in Fig. 1 overlaid on a single annual timebase. In Fig. 2, the long-term monthly mean values of CCN concentration at the extremes of the five available supersaturations⁴ are plotted against the long-term monthly mean aerosol MSA concentrations. These plots show clearly that for variations in MSA concentration forced by seasonal factors, there is a relationship between CCN and MSA, but it is not simple. There is an apparently significant intercept at zero MSA, and an indication that the gradient of CCN concentration against MSA concentration flattens at high MSA concentrations. This effect is very clear for the individual monthly values of total particle concentration⁷ plotted in the lowest panel of Fig. 2.

The positive intercepts that can be inferred from Fig. 2 at zero MSA are ~40 cm⁻³ for CCN active at 0.23%, 60 cm⁻³ for CCN active at 1.2% and 80 cm⁻³ for total particles, the last corresponding to particles active at a very large supersaturation (>200%). This consistent picture implies the existence of perhaps 40 cm⁻³ of CCN active at 0.23% that contain little or no MSA, one possible explanation for which could be that a significant fraction of CCN is not derived from DMS. Sea salt must provide one non-DMS source for CCN: Cape Grim is located in the 'roaring forties' where strong winds are common and sea salt is a principal contributor to aerosol mass.⁶ Previous measurements at Cape Grim⁸ suggest mean sea-salt number concentrations of ~10 cm⁻³, which will rise in extreme wind

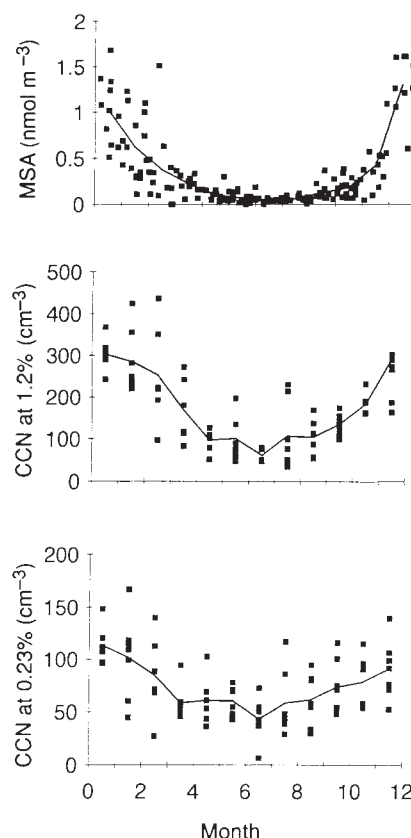


FIG. 1 Climatological mean (1981-1989) concentrations of aerosol MSA and CCN active at 0.23% and 1.2% supersaturation (solid lines). The data from which the means were calculated are shown as filled squares. In the case of MSA these data came from monthly samples before 1985, weekly samples thereafter^{6,7}. The CCN data used are the monthly medians reported by Gras⁴.

conditions. But a considerable fraction of the ~40 cm⁻³ must still originate from other sources, which at this stage remain unidentified. On the basis of aerosol chemistry measurements, we previously suggested a source of aerosol sulphate at Cape Grim additional to that of DMS oxidation³.

The roll-off in nucleus concentration for large MSA values (Fig. 2) is consistent with nonlinearities inherent in aerosol formation and growth by nucleation-condensation of supersaturated acid vapour. Once an aerosol has developed sufficient surface area to accommodate the flux of condensing acid molecules, new particle formation ceases. At the same time the rate of increase in particle size slows as average size increases^{9,10}. The limit on nucleation is clear for the total particle data in Fig. 2. With increasing MSA, particles continue to grow to sizes sufficient to be active as CCN, but the expected slowing of growth is evident in Fig. 2.

We used a numerical model of droplet nucleation and growth¹¹, incorporating observed mean summer and winter CCN supersaturation spectra from Cape Grim, to assess the maximum supersaturation likely for stratiform cloud of the type central to the hypothesis of Charlson *et al.*¹. For updraft velocities of 10, 20 and 40 cm s⁻¹, the model predicts maximum supersaturations in summer/winter of 0.10/0.15% at 10 cm s⁻¹, 0.16/0.26% at 20 cm s⁻¹ and 0.27/0.45% at 40 cm s⁻¹, suggesting that our measurements at 0.23% are most relevant to the hypothesis of ref. 1. From the plot in Fig. 2 for 0.23% supersaturation, we conclude that MSA may be relatively unimportant as a component of CCN in winter (MSA approaches zero). Its role would clearly be larger in summer, but given the roll-off at higher MSA and the possibility that a significant number of

CCN not derived from DMS might also exist in this season, the relationship between CCN and aerosol MSA concentrations would remain much weaker than a linear relationship.

Given these conclusions, it is essential to consider the extent to which MSA can be used as a surrogate for the DMS from which it is derived. A recent report¹² for several Northern Hemisphere sites and different seasons between 1986 and 1989 gave the DMS/MSA ratios for six separate field studies as 8.2, 4.6, 7.6, 8.5, 8.2 and 4.8. The consistency of these values certainly suggests that MSA may be a useful tracer for DMS. Published data on DMS from Cape Grim³ cover only a 20-month period, and have some gaps, but plotted against the corresponding aerosol MSA values yield the relationship shown in Fig. 3. Some scatter is expected, as the DMS values plotted result on average from only three 1-hour samples within a given week, whereas MSA values derived from aerosol samples were integrated over the whole week. About a third of the points arise from a single DMS measurement. But the relationship is significant and indicates a relatively consistent molar ratio of about 8 for DMS/MSA throughout the year, a value very consistent with those reported in ref. 12.

The existence of a strong temperature dependence in the relative efficiency with which DMS is oxidized to MSA, compared with oxidation to SO_2 (the so-called branching ratio¹³), does not mitigate against a stable DMS/MSA ratio at Cape Grim because annual temperature variation is small. Long-term records⁶ show a mean summer–winter temperature range of only 287.5 to 282.7 K, leading to a predicted change of only 38% in branching ratio¹³ in the direction which would produce a decrease in DMS/MSA ratio from summer to winter. This is a small perturbation relative to DMS and MSA concentration

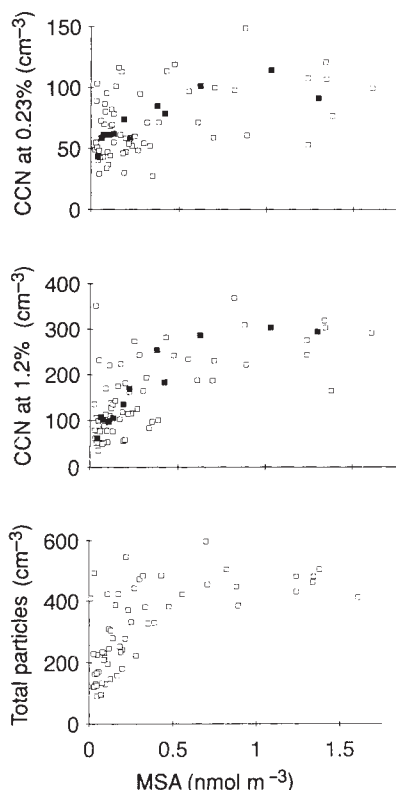


FIG. 2 Variation of mean (1981–1989) monthly CCN concentrations (filled squares) at 0.23% and 1.2% supersaturation as a function of mean monthly MSA concentration at Cape Grim (upper two plots). The individual monthly data points are also shown as open squares. Lowest plot: monthly mean (1981–1989) total particle concentration as a function of MSA concentration.

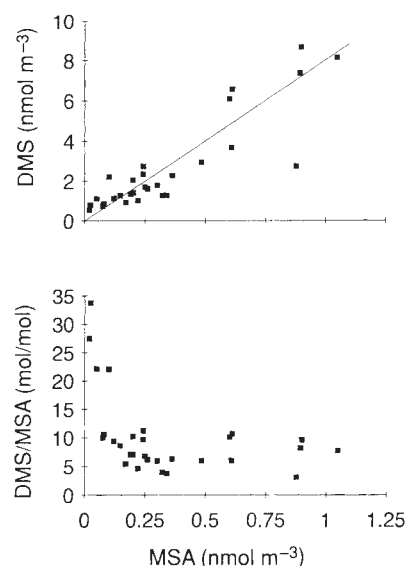


FIG. 3 Upper: DMS concentration against aerosol MSA concentration at Cape Grim, October 1988–May 1990. The line shown has a slope of 8. Lower: DMS/MSA molar ratio derived from the same data, plotted against MSA concentration.

variations of more than an order of magnitude between summer and winter. From the lower panel in Fig. 3, in which data from the upper panel are replotted as DMS/MSA ratio against MSA concentration, it is clear that there is no decrease in the DMS/MSA ratio in winter (when MSA is low). If anything the opposite is true, although caution is needed in interpreting the few high values, as these arise from the ratio of single 1-hour DMS samples to integrated weekly MSA samples, and the uncertainty for these points is therefore large.

We believe there are reasonable grounds for assuming that the measured CCN–MSA relationship at Cape Grim is a useful indicator of the CCN–DMS relationship central to the hypothesis of Charlson *et al.*¹ We stress, however, that our conclusions should at this stage be seen as specific to Cape Grim. Without confirmatory data from elsewhere we cannot exclude the possibility, for example, that the unidentified source of CCN at very low MSA concentrations might be a regional, rather than hemispheric, phenomenon. Nevertheless, our results indicate the need for concurrent measurement of CCN, MSA and DMS concentration at other remote marine sites to investigate, first, the relative importance of CCN not derived from DMS, and second, the possibility of a lowering in the efficiency of CCN production from DMS as CCN concentration increases. □

Received 30 May; accepted 10 September 1991.

1. Charlson, R. J., Lovelock, J. E., Andreae, M. O. & Warren, S. G. *Nature* **326**, 655–661 (1987).
2. Bates, T. S. *et al. Global biogeochem. Cycles* **3**, 299–304 (1989).
3. Ayers, B. P., Ivey, J. P. & Gillett, R. W. *Nature* **349**, 404–406 (1991).
4. Gras, J. L. *Geophys. Res. Lett.* **17**, 1565–1567 (1990).
5. Ayers, G. P., Ivey, J. P. & Goodman, H. S. *J. Atmos. Chem.* **4**, 173–185 (1986).
6. Baseline Series, Ann. Rep. Baseline Atmospheric Program (Australia) (Bureau of Meteorology, Melbourne 3000, Australia, 1976–1988).
7. Gras, J. L. *J. Atmos. Chem.* **11**, 89–106 (1990).
8. Gras, J. L. & Ayers, G. P. *J. geophys. Res.* **88**, 10661–10666 (1983).
9. McMurray, P. H. & Friedlander, S. K. *J. Colloid Interface Sci.* **64**, 248–257 (1978); *Atmos. Envir.* **13**, 1635–1651 (1979).
10. McMurray, P. H. *J. Colloid Interface Sci.* **78**, 513–527 (1980).
11. Ayers, G. P. & Larson, T. V. *J. Atmos. Chem.* **11**, 143–167 (1990).
12. Burgermeister, S. & Georgii, H.-W. *Atmos. Envir.* **A25**, 587–595 (1991).
13. Hynes, A. J. *et al. J. Phys. Chem.* **90**, 4148–4150 (1986).

ACKNOWLEDGEMENTS. We thank the staff at the Cape Grim Baseline Air Pollution Station for their work over fifteen years. This study was supported by the Cape Grim research fund, administered jointly by the Australian Bureau of Meteorology and CSIRO.

Freshwater flux forcing of decadal and interdecadal oceanic variability

Andrew J. Weaver*, E. S. Sarachik†
& Jochem Marotzke‡

* Department of Meteorology and Centre for Climate and Global Change Research, McGill University, 805 Sherbrooke Street West, Montreal, P.Q., H3A 2K6, Canada

† Joint Institute for the Study of the Atmosphere and Ocean, University of Washington, AK-40, Seattle, Washington, 98195, USA

‡ Department of Earth, Atmospheric and Planetary Sciences, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA

THE ocean's thermohaline circulation, driven by fluxes of freshwater and heat through the ocean's surface, is important in the transport of heat from low to high latitudes. Changes in the intensity of this circulation, and hence the poleward heat transport, would have a significant effect on global climate. Numerous observations of the air-sea-ice climate system in and around the North Atlantic show significant variability on decadal and interdecadal timescales¹⁻¹¹; these timescales suggest that the source of the variability lies in the ocean itself. Here we present the results of three numerical experiments which show the importance of freshwater flux forcing (the difference between precipitation and evaporation, or ' $P-E$ ') in exciting decadal and interdecadal oceanic variability. If a sufficiently strong local minimum exists in this $P-E$ forcing field (as is observed over the Greenland Sea), self-sustained oscillations of the ocean circulation may be excited. We propose that such variability may be important in interpreting observations of decadal/interdecadal variability in the air-sea-ice climate system.

In many ocean modelling studies, researchers use *restoring boundary conditions* on both surface temperature (T) and salinity (S). These boundary conditions imply a direct feedback between oceanic (T_O, S_O) surface values and atmospheric (T_A, S_A) forcing variables, whereby the flux of heat (Q_T) and salt (Q_S) out of the ocean is given by the simple newtonian damping law $Q_T = C_T(T_O - T_A)$, $Q_S = C_S(S_O - S_A)$ for heat and salt, respectively. Here C_T and C_S are positive constants usually taken to be equal. Although there is substantial evidence justifying the use of the restoring boundary condition on temperature¹², there is no physical justification for its use on salinity. A more appropriate set of boundary conditions are *mixed boundary conditions* in which the salinity flux is fixed but the restoring boundary condition on temperature is retained.

The use of mixed boundary conditions in modelling the thermohaline circulation (THC) has produced phenomena with important implications for the predictability of both past and present climate, and climate change. For example, multiple equilibria have been found and the transition between these equilibria studied¹³⁻¹⁸. Apparently contradictory results were observed in the transient behaviour of the THC during long-time integrations of the Bryan-Cox ocean general circulation model (OGCM). In some studies^{14,15}, decadal and interdecadal internal variability was prevalent in the THC, whereas in others^{13,16-18} no such variability was found (one of the two-basin experiments described in ref. 18 indicated the presence of decadal/interdecadal oscillations in the region of the Antarctic circumpolar current). Here we reconcile these results by examining the importance of the freshwater flux boundary condition.

We present the results of three numerical experiments conducted using the primitive equation OGCM¹⁹. The model geometry consists of a flat-bottomed basin 60° wide, extending from 0° to 64° N and forced by steady winds. For computational economy we use coarse horizontal resolution (4° meridional $\times$ 3.75° zonal)

with 15 vertical levels in which the grid-box thickness varies from 50 m near the surface to 500 m near the bottom (at 4,500 m). The details of the wind forcing and diffusion and viscosity parameters are given elsewhere^{16,17}. As in refs 16 and 17, we use linearized horizontal momentum equations. In experiment 1, hereinafter WSBC, we start the model from rest using zonally uniform restoring boundary conditions at the surface obtained from global averages of surface data²⁰. In experiments 2 and 3 (hereinafter CBC1 and CBC2, respectively), we run the same model under idealized zonally uniform restoring boundary conditions on both temperature and salinity. For the surface restoring temperatures used in WSBC, see ref. 21; for CBC1 and CBC2 (where the restoring temperatures are identical), see ref. 17. All three of these profiles consist of an equatorial restoring temperature of $T_A = 27^\circ\text{C}$ and a smooth transition to $T_A = 0^\circ\text{C}$ at 64° N. Figure 1a shows the zonally uniform surface restoring salinity for each of the three experiments, and Fig. 1b shows the net (salinity and temperature) surface restoring density.

The three salinity profiles represent three different situations: salinity dominant, WSBC; temperature dominant, CBC1; salinity important, CBC2. We adopted this terminology to indicate the relative importance of freshwater flux against thermal forcing, as deduced by the strength of the meridional gradients shown in Fig. 1a and c. The difference between the WSBC and CBC1/CBC2 experiments is therefore given by the strength of the $P-E$ forcing. The WSBC boundary conditions were derived from global surface data, whereas the CBC1 and CBC2 boundary conditions were obtained from idealized fields. Below we

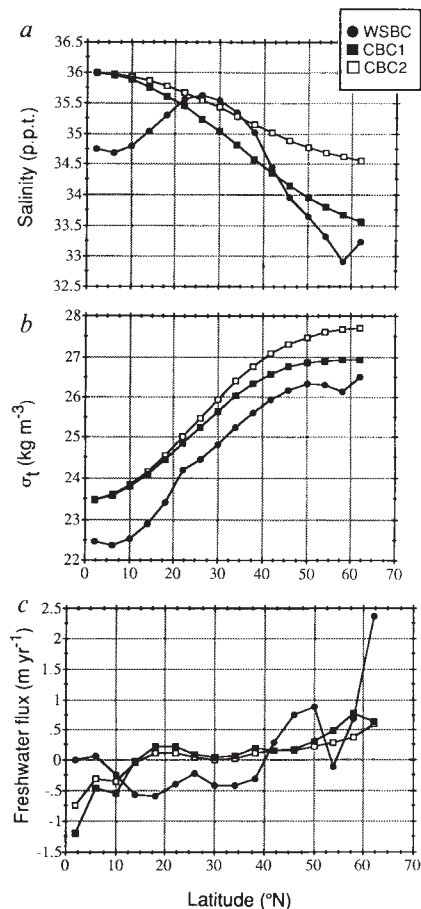


FIG. 1 a, Restoring boundary condition on salinity (parts per thousand) used in the three experiments WSBC, CBC1 and CBC2. b, Density, σ_t , as defined by the restoring temperatures and salinities. c, Zonally averaged freshwater flux in (m yr^{-1}) as diagnosed from the steady states of the WSBC, CBC1 and CBC2 experiments under restoring boundary conditions.

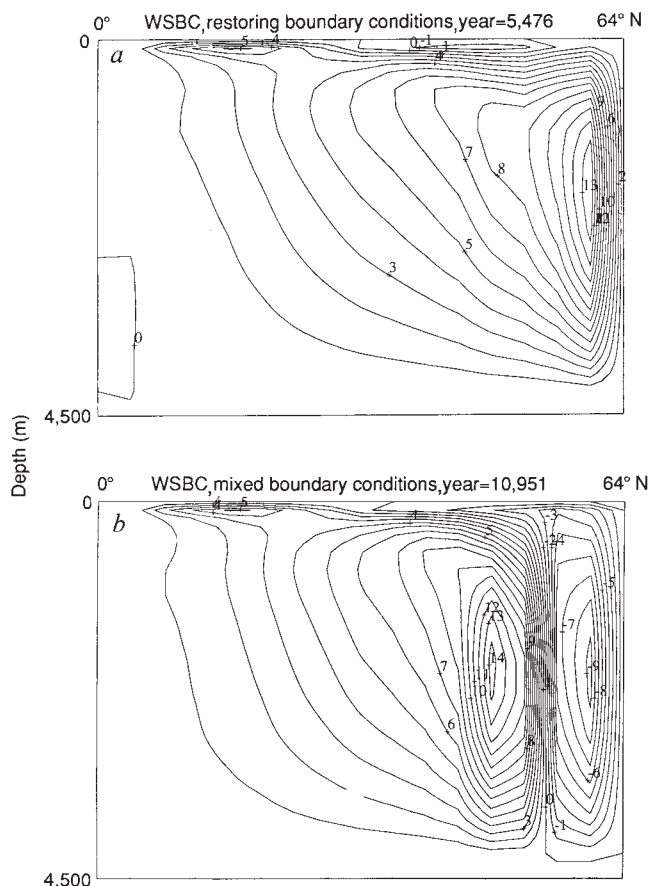


FIG. 2 Steady states obtained from the WSBC experiment. *a*, Under restoring boundary conditions (after 5,476 years of integration); *b*, under mixed boundary conditions (after a further integration of 10,951 years from the steady state of *a*). Contours of meridional overturning in Sv are shown. The steady state shown in (*a*) is unstable under mixed boundary conditions, whereas the steady state in (*b*) is stable.

shall use the results of the WSBC experiment for a comparison with the $P-E$ forcing of the North Atlantic.

Figure 2*a* illustrates the steady state obtained after 5,476 years of integration of WSBC under restoring boundary conditions. In this steady state ~ 13 Sv of deep water is formed at the northern wall, where the surface restoring density is largest (Fig. 1*b*). We instantaneously diagnosed the freshwater flux required to maintain this steady state¹³ and switched the boundary condition on salinity from a restoring boundary condition to a time-invariant, specified freshwater flux. If the steady state obtained under restoring boundary conditions were stable to infinitesimal perturbations under the consistent mixed boundary conditions, we would expect no transition to occur after the switch. In WSBC, the steady state obtained under restoring boundary conditions was unstable upon a switch to mixed boundary conditions. We eventually reached an equilibrium (portrayed in Fig. 2*b* at year 10,951 of further integration); but this steady state was fundamentally different from the one portrayed in Fig. 2*a*. Deep water no longer formed at the northern wall but rather at 54° N, the local minimum in $P-E$ (Fig. 1*c*). A reverse cell is now present at high latitudes.

The transition to equilibrium in WSBC is interesting as it is apparently chaotic (Fig. 3*a*), with significant energy on a decadal timescale (Fig. 3*b*). The existence of the variability is linked to the local evaporative region near 54° N (Fig. 1*c*). When the THC is weak, it slowly passes through this evaporative region and hence the surface waters become more saline. This induces

sinking at 54° N (and the subsequent generation of the reverse cell seen in Fig. 2*b*), which in turn causes the THC to intensify. The intensified THC passes rapidly through the evaporative region and hence the surface waters do not become as saline. Deep water then forms at high latitudes (Fig. 2*a*) until high-latitude freshening dominates (Fig. 1*c*) and the THC slows down. This whole process repeats itself. The timescale of the oscillation is determined by the time taken for salinity and temperature anomalies, formed in the local evaporative region, to be advected to the northern boundary¹⁴. The oscillation has a profound effect on the northward transport of heat by the ocean and hence on the net basin-averaged surface heat flux, especially at high latitudes. Figure 3*a* shows the basin-averaged heat flux throughout the integration of WSBC. Variations of between -10 and 6 W m^{-2} occur on a decadal timescale. Because most of this variability occurs at high latitudes, the heat flux variations are even stronger there. Such a marked variability in the ocean-to-atmosphere heat flux over the relatively short (decadal) timescale could have a considerable impact on both local and global climate.

The steady state obtained under restoring boundary conditions (Fig. 4*a*) in CBC1 (where salinity forcing has a minor role) was stable on the switch to mixed boundary conditions (confirmed by an integration of 1,369 years under mixed boundary conditions). In CBC1 we found no internal variability and

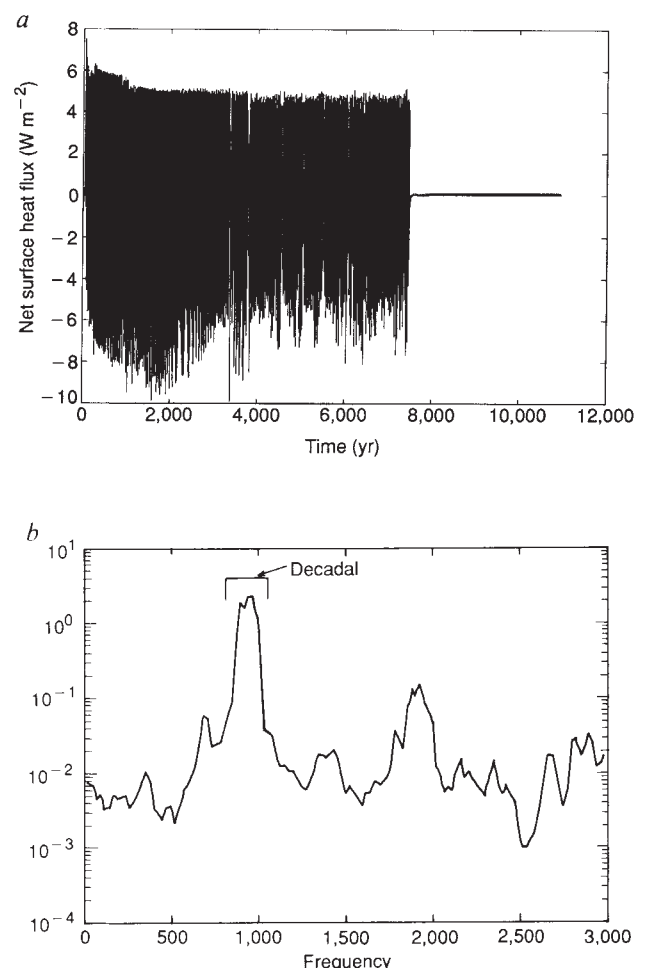


FIG. 3 *a*, Basin-averaged net surface heat flux (W m^{-2}) for the WSBC integration under mixed boundary conditions. *b*, Power spectral density (using a 256-point fast Fourier transform) of the surface heat flux shown in (*a*) for the first 8,214 years of integration of the WSBC experiment. The frequency (*x*-axis) corresponds to the number of cycles over the 8,214 years of integration.

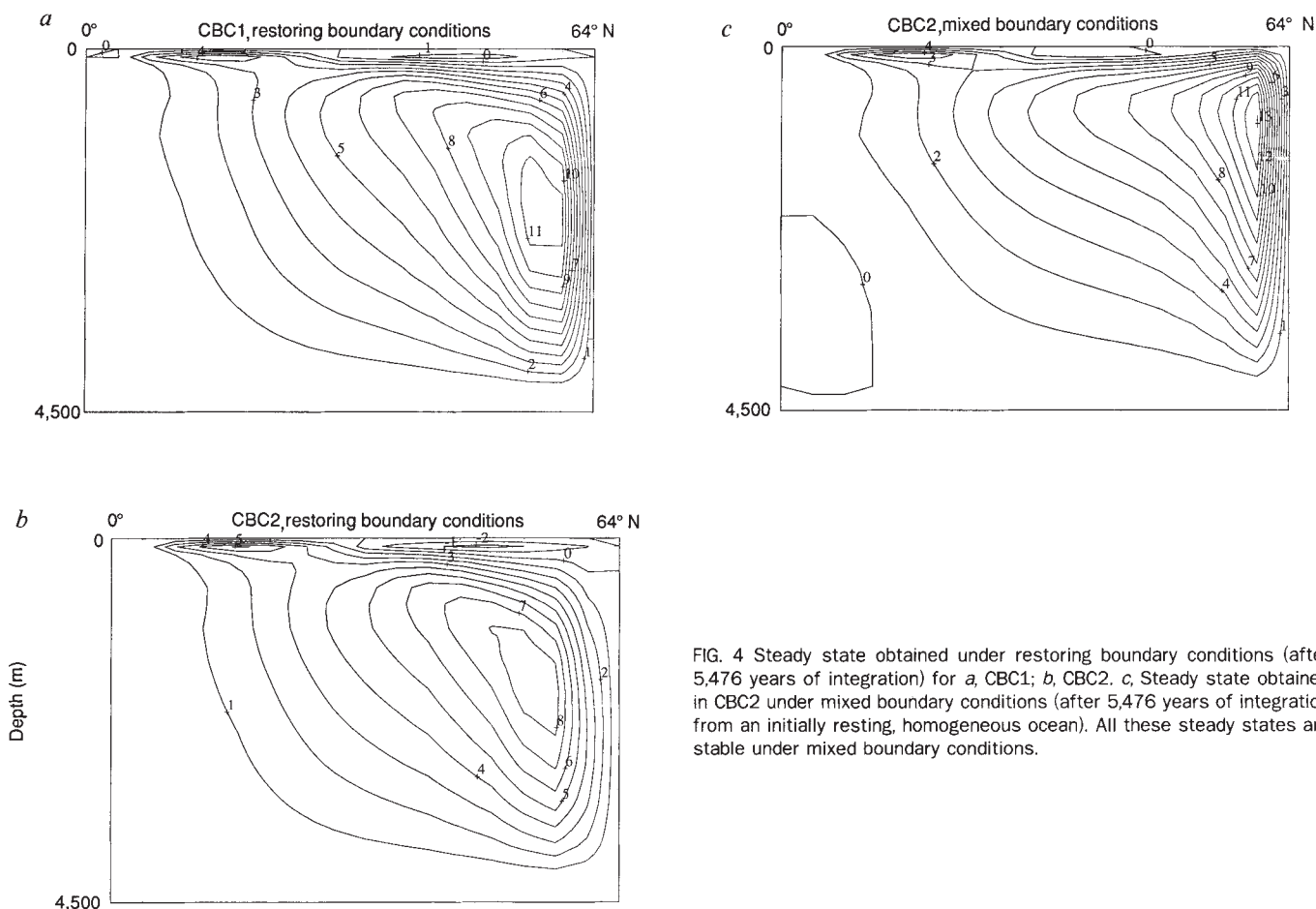


FIG. 4 Steady state obtained under restoring boundary conditions (after 5,476 years of integration) for *a*, CBC1; *b*, CBC2. *c*, Steady state obtained in CBC2 under mixed boundary conditions (after 5,476 years of integration from an initially resting, homogeneous ocean). All these steady states are stable under mixed boundary conditions.

only one steady state, corresponding to the steady state obtained under restoring boundary conditions. This was further evident from an experiment in which we took an initially resting ocean with constant temperature and salinity everywhere (equal to the basin average at the steady state of CBC1 under restoring boundary conditions) and started the model using mixed boundary conditions from the onset. The steady state that we obtained was identical to that shown in Fig. 4*a*, with the transition to this state showing no decadal/interdecadal variability. This is very different from the results of a WSBC experiment where we spun up an initially homogeneous, resting ocean under mixed boundary conditions from the onset. In the approach to equilibrium the system showed internal variability, with dominant signal at decadal timescale (Fig. 3*b*), for thousands of years.

In our third situation (CBC2), the steady state obtained under restoring boundary conditions (Fig. 4*b*) was again stable on switching to mixed boundary conditions (confirmed by an integration of 1,369 years under mixed boundary conditions). We now found two equilibria, however, both of which were stable. This was confirmed by an integration from rest using mixed boundary conditions from the onset, as discussed above. The resulting steady state (Fig. 4*c*) is significantly different from the one obtained under restoring boundary conditions, although no internal variability was found in the spin-up, as there were no significantly strong local minima in the field of surface freshwater flux (Fig. 1*c*). The equilibrium shown in Fig. 4*b*, although stable to infinitesimal perturbations, could be made unstable by adding small perturbations to the high-latitude surface salinity (as in ref. 13). The equilibrium found when the THC was smoothly reestablished was identical to Fig. 4*c*¹⁶.

As discussed earlier, there has been considerable evidence for decadal and interdecadal variability in and around the North Atlantic. Observed estimates of annual mean $P - E$ over the

North Atlantic^{22,23} show that over the subtropics there is net evaporation with net precipitation over mid-latitudes (as in WSBC; Fig. 1*c*). In the region of the Greenland Sea, where deep water forms³, there is a local net evaporation maximum (compare with Fig. 1*c*) with net precipitation at higher latitudes (where runoff also becomes important⁷). The $P - E$ pattern of the North Atlantic thus resembles that used in WSBC, and so we might expect internal decadal and interdecadal variability in the THC of the ocean there. □

Received 1 May; accepted 10 September 1991.

1. Ghil, M. & Vautard, R. *Nature* **350**, 324–327 (1991).
2. Gray, W. M. *Science* **249**, 1251–1256 (1990).
3. Schlosser, P., Bönsch, G., Rhein, M. & Bayer, R. *Science* **51**, 1054–1056 (1991).
4. Lazier, J. R. N. *Atmos. Ocean* **18**, 227–238 (1980).
5. Dickson, R. R., Meincke, J., Malmberg, S.-A. & Lee, A. J. *Prog. Oceanogr.* **20**, 103–151 (1988).
6. Mysak, L. A. & Manak, D. K. *Atmos. Ocean* **27**, 376–405 (1989).
7. Mysak, L. A., Manak, D. K. & Marsden, R. F. *Clim. Dynam.* **5**, 111–113 (1990).
8. Gloersen, P. & Campbell, W. J. *Nature* **352**, 33–36 (1991).
9. Cattle, H. *Polar Rec.* **22**, 485–498 (1985).
10. Ikeda, M. *Atmos. Ocean* **28**, 106–139 (1990).
11. Krishnamurti, T. N., Chu, S.-H. & Iglesias, W. *Arch. Met. Geoph. Bioclim., Ser. A* **34**, 385–425 (1986).
12. Haney, R. L. *J. phys. Oceanogr.* **1**, 241–248 (1971).
13. Bryan, F. *Nature* **323**, 301–304 (1986).
14. Weaver, A. J. & Sarachik, E. S. *Atmos. Ocean* **29**, 197–231 (1991).
15. Weaver, A. J. & Sarachik, E. S. *J. phys. Oceanogr.* **21**, 1470–1493 (1991).
16. Marotzke, J. *J. phys. Oceanogr.* **21**, 903–907 (1991).
17. Marotzke, J. & Willebrand, J. *J. phys. Oceanogr.* **21**, 1372–1385 (1991).
18. Marotzke, J. thesis, Ber. Int. Meeresk., Kiel (1990).
19. Cox, M. D. *Tech. Rep. No. 1* (GFDL Ocean Group, Princeton University, New Jersey, 1984).
20. Levitus, S. *NOAA Prof. Pap. 13* (US Department of Commerce, NOAA, Rockville, Maryland, 1982).
21. Weaver, A. J. & Sarachik, E. S. *J. phys. Oceanogr.* **20**, 600–609 (1989).
22. Baumgartner, A. & Reichel, E. *The World Water Balance* (Elsevier, New York, 1975).
23. Schmitt, R. W., Bogden, P. S. & Dorman, C. E. *J. phys. Oceanogr.* **19**, 1208–1221 (1989).

ACKNOWLEDGEMENTS. We thank C. Lin, L. Mysak, S. Power and M. Winton for discussions. This research was supported by NSERC and FCAR funding (A.J.W.) and by a grant from the NOAA Office of Climate and Atmospheric Research to E.S. at the University of Washington Experimental Climate Forecast Center. J.M. was supported through NSF.

Geochemical constraints on source region of Cretaceous/Tertiary impact glasses

H. Sigurdsson^{*†}, Ph. Bonté^{*}, L. Turpin[‡],
M. Chaussidon[§], N. Metrich^{||}, M. Steinberg[¶],
Ph. Pradel[¶] & S. D'Hondt[†]

^{*} Centre des Faibles Radioactivités, CNRS-CEA, 91198 Gif-sur-Yvette, France

[†] Graduate School of Oceanography, University of Rhode Island, Narragansett, Rhode Island 02882, USA

[‡] Commissariat à l'Energie Atomique, Centre de la Vallée du Rhône, 30205 Bagnols-sur-Cèze, France

[§] CRPG-CNRS, 54501 Vandœuvre les Nancy, France

^{||} Laboratoire Pierre Süe, CEN-Saclay 91191 Gif-sur-Yvette, France

[¶] Laboratoire de Géochimie Sédimentologie, Université d'Orsay, 91405 France

THE 50-cm-thick Cretaceous/Tertiary boundary layer at Beloc in Haiti contains spherules of silicic black glass, 1–8 mm in diameter, which have been attributed to impact fusion of continental crust¹. Calcic yellow glass (up to 30 wt% CaO) forms a coating on and streaks within some black glass spherules, and has been attributed to melting of carbonate-rich sediment¹, or organic-rich or pyritic limestone². Here we present new trace-element and stable and radiogenic isotope data which show that the silicic glass is derived from continental crust of andesitic composition, whereas the high-Ca glass formed by melting of evaporite-rich sediment. This is confirmed by melting experiments with evaporite and andesite at 1,200–1,400 °C, which approximately reproduce the high-Ca glass. The temperature-dependent variation of sulphur content in synthetic high-Ca glasses indicates a formation temperature of 1,300 °C for the Haiti glasses. The geology of the impact site inferred from the geochemistry of the Haiti glasses matches the lithologies found in the 180-km Chicxulub structure, which occurs in Cretaceous evaporite deposits in Mexico. The high sulphur content of the calcic glasses suggests that the impact may have generated significant emissions of sulphur dioxide to the atmosphere, causing short-term global cooling.

The Haiti glasses have a much lower volatile content than volcanic glasses³ and an H₂O/CO₂ ratio consistent with a high partial pressure of carbon dioxide in the confining gas phase². Unlike young tektites, however, these Cretaceous/Tertiary boundary (K/T) glasses are surrounded by a smectite alteration

shell, and they also lack the lechatelierite found in most tektites. Geochemical study of the Haiti impact glass particles shows that the black and yellow glasses are relatively homogeneous, with standard deviation in trace element content from 8 to 15% (Table 1). The black glass is depleted in alkalis, U, Th and Ta relative to average upper continental crust⁴. It has similarities to Archean crust and mudstone (Fig. 1), such as low La/Sc and Th/Sc ratios, low U and Th, low ratio of light to heavy rare earth elements (REE) and lack of Eu/Eu* anomaly⁴. But the glass is considerably depleted in Cs, Cr, Ni and Co compared with Archean crust, and thus this terrane is an unlikely source. The glass is most closely comparable to andesite from continental margins¹ and is virtually indistinguishable from some Mexican andesites (Fig. 1). The high ⁸⁷Sr/⁸⁶Sr ratio of 0.70901 in the black glass is indicative of a continental crustal source, or an igneous source contaminated by highly radiogenic continental sediment, and several andesites have ⁸⁷Sr/⁸⁶Sr values in this range⁵. The δ¹⁸O value of the siliceous black Haiti glass is also consistent with this source (Table 1).

The yellow glass is depleted by 10–20% relative to black glass in rare earths and most other trace elements of small ionic radius, and in most major elements (Table 1). It is enriched, however, in Ca and Mg. If the yellow glass is the product of a mixture of calcic sediment (evaporite or carbonate) and silicate melt similar to the black glass, then dilution of REE and other trace elements absent in the calcic sediment is expected because of added CaO and MgO. This is consistent with the observed REE dilution factor of 10–20% which corresponds to the addition of 18 wt% CaO and MgO to the yellow glass¹. The alkali and alkaline earth elements (Na, K, Rb, Cs) are depleted in the yellow glass (Fig. 2). Their depletion is attributed to volatilization during impact melting, as shown by melting experiments reported below.

The sulphur content (up to 1%) in the yellow glass is higher than generally observed in carbonate rocks⁶. Furthermore, the yellow glass lacks the enrichment in uranium or chalcophile elements generally associated with pyritic or organic-rich limestones. The high oxygen fugacity³ and high δ³⁴S value of 13.2‰ in the yellow glass (Table 1) show that sulphur was present as sulphate in the source. This and the strontium data discussed below indicate that the sediment affected by impact included gypsum or anhydrite-bearing evaporite deposits. If this is sulphate of entirely marine origin, the isotope value indicates an evaporate of lower Cretaceous age⁷. Because of possible mixing of sources and possible non-marine character of the evaporite, however, the sulphur isotope data cannot constrain the age. The δ¹⁸O value of 13.3‰ in the high-Ca yellow glass is also comparable to that of Cretaceous evaporites^{7,8}, and the δ¹¹B (Table 1) of yellow glass is within the range of borates from non-marine

FIG. 1 Trace and major element distributions in the black Haiti impact glass are closely similar to those in Mexican andesite³³. Another rock type with comparable chemistry is Archean mudstone⁴, but Cs, Cr, Co and Ni anomalies exclude this terrane as source of the impact glass.

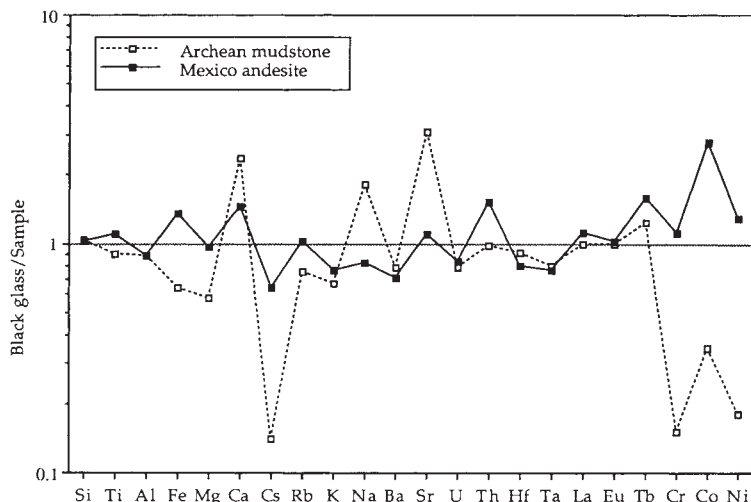


TABLE 1 Chemical composition of K/T boundary impact glasses and synthetic glasses

	1	2	3	4	5	6	7	8
SiO ₂	64.29 (2.1)	57.93 (0.95)	57.46 (1.1)	46.76 (1.7)	43.32 (2.2)	45.82 (0.94)	45.78 (1.2)	41.13 (1.9)
TiO ₂	0.65 (0.07)	1.06 (0.19)	0.94 (0.08)	0.61 (0.07)	0.82 (0.07)	0.82 (0.07)	0.97 (0.15)	0.76 (0.11)
Al ₂ O ₃	15.21 (0.30)	16.38 (0.35)	16.43 (0.44)	12.50 (0.81)	13.01 (1.1)	11.92 (0.26)	13.32 (0.69)	11.61 (0.26)
Fe ₂ O ₃	5.32 (0.38)	6.24 (0.36)	6.73 (0.55)	4.91 (0.21)	5.43 (0.50)	5.30 (0.11)	6.20 (0.76)	5.30 (0.24)
MgO	2.67 (0.21)	4.00 (0.29)	4.05 (0.43)	4.27 (0.15)	3.20 (0.37)	3.08 (0.21)	3.65 (0.74)	2.99 (0.22)
CaO	6.52 (1.2)	6.52 (0.28)	6.97 (1.1)	27.00 (2.7)	29.26 (5.5)	28.50 (1.4)	23.41 (4.0)	33.29 (1.7)
Na ₂ O	3.56 (0.29)	4.37 (0.29)	4.55 (0.35)	2.36 (0.30)	2.94 (0.34)	2.29 (0.16)	1.81 (0.97)	1.00 (0.20)
K ₂ O	1.58 (0.12)	2.27 (0.08)	2.44 (0.58)	0.62 (0.12)	1.34 (0.13)	1.21 (0.13)	0.82 (0.11)	0.30 (0.05)
SO ₃	0.10 (0.05)	0.14 (0.07)	0.14 (0.10)	0.83 (0.05)	0.34 (0.17)	0.78 (0.18)	1.33 (0.52)	3.34 (0.88)
Cs (p.p.m.)	0.86			0.23				
Rb	45			17				
Ba	450			550				
Sr	550			1620				
U	1.25			1.25				
Th	6.2			5.2				
Hf	3.2			3.0				
Ta	0.36			0.35				
La	20			18				
Ce	43			37				
Nd	21			18				
Sm	4.5			4.2				
Eu	1.20			1.05				
Tb	0.7			0.67				
Yb	2.7			2.3				
Lu	0.39			0.37				
Cr	30			25				
Co	14			11				
Ni	18			13				
Sc	19			17				
Zr	110			95				
Zn	57			50				
F*	23			188				
B†	2			28				
Li†	7.4			12.5				
Be†	1.4			1.1				
δ ³⁴ S†	not done			+13.2‰				
δ ¹¹ B†	+3.4‰			-2.1‰				
δ ¹⁸ O	+7.7‰			+13.3‰				
⁸⁷ Sr/ ⁸⁶ Sr	0.70901			0.70799				

* Determined by nuclear microprobe, PIGME³¹.

† Determined with Cameca IMS-3F ion microprobe at CRPG, Nancy; CDT standard for sulphur and NBS-951 standard for boron; standard error $\pm 10\%$ in concentration and $\pm 4\%$ in isotope ratio. The oxygen isotope value $\delta^{18}\text{O}$ is reported with respect to SMOW ($\delta^{18}\text{O} = [(^{18}\text{O}/^{16}\text{O})_{\text{sample}} / (^{18}\text{O}/^{16}\text{O})_{\text{SMOW}}] - 1$, given in ‰). The Sr contents of black and yellow glass samples analysed for ⁸⁷Sr/⁸⁶Sr ratio are 330 and 1,600 p.p.m., respectively. The ⁸⁷Sr/⁸⁶Sr values are corrected for ⁸⁷Rb decay since the time of the K/T boundary, and the analytical error is ± 0.00002 at the 95% confidence limit³². Other trace elements were determined by neutron activation analysis. The accuracy (2σ) is better than 5% for Th, Hf, La, Ce, Sm, Eu, Co and Sc; better than 10% for Rb, U, Ta, Nd, Tb, Yb, Lu, Zr and Zn; better than 20% for Cs, Ba, Sr, Cr and Ni.

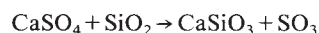
Column 1: Average composition of black impact glass, Haiti; mean and standard deviation of 74 electron microprobe analyses of major elements (wt%). Neutron activation analysis of trace elements, in p.p.m.; average of eight spherules. 2: Synthetic high-silica glass from fusion of andesite in presence of gypsum; mean of five analyses. Andesite from Pico de Orizaba volcano, Mexico (PU-170¹³). 3: Synthetic high-silica glass from fusion of Mexican andesite in presence of calcite; mean of 11 analyses. 4: Average composition of yellow high-Ca impact glass, Haiti; mean of 13 electron microprobe analyses of major elements. Trace elements represent average neutron activation analysis of eight particles (p.p.m.). 5: Composition of synthetic yellow glass from fusion of evaporite and andesite at 1,402 °C. Average of four analyses. Evaporite was gypsum from Messinian evaporite, Calabria, Italy. 6: Composition of synthetic yellow glass from fusion of evaporite and andesite at 1,308 °C. Average of three analyses. 7: Composition of synthetic yellow glass from fusion of evaporite and andesite at 1,208 °C. Average of three analyses. 8: Composition of synthetic yellow glass from fusion of evaporite and andesite at 1,167 °C. Average of four analyses.

evaporites⁹. The bulk sediment $\delta^{11}\text{B}$ value of the yellow glass cannot, however, be strictly compared to borate isotope data.

Strontium enrichment in yellow glass is attributed to a contribution from the calcic sediment, where strontium substitutes for calcium. The strontium content is higher than is typical in open ocean carbonate sediments¹⁰. For example, the Sr/Ca atomic ratio in open ocean calcite is 0.001–0.0014 (ref. 11), compared with 0.0046 in the yellow glass. The high strontium is compatible with a shallow-water aragonite source, or an evaporite source, as the Sr content can be high in gypsum or anhydrite¹², and evaporites frequently contain celestite. The Sr isotope ratio of the yellow glass (⁸⁷Sr/⁸⁶Sr = 0.70799) is much lower than that of the black glass and is consistent with an evaporite origin.

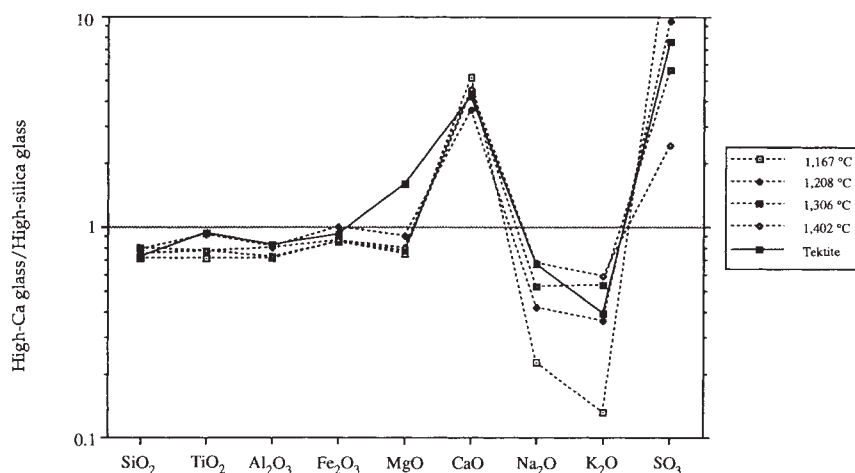
We have reproduced the high-Ca yellow glass in the laboratory

by melting evaporite in contact with a melt of continental crust or andesite composition. The melting of the Ca-rich sediment component (carbonate or sulphate-rich evaporite) leads to desilication of the co-existing high-silica melt, with evolution of the volatile phase during reactions of the type:



Ten melting experiments were carried out in platinum tubes, 3 mm in diameter, in a vertical furnace at 1 atm and 1,047–1,406 °C for one and two hours. Initial compositions were powders of gypsum, calcite and andesite¹³. About 60 mg of the andesite powder was overlain by the same amount of the Ca-rich powder. Yellow, high-Ca glass was produced by quenching in all experiments, coexisting with high-silica glass with composi-

FIG. 2 Oxide ratios of coexisting high-silica and high-Ca glasses in Haiti K/T boundary spherules and in 1-atm fusion experiments with evaporite and andesite, showing identical compositional relations, including alkali loss from the high-Ca glass due to volatilization. Experiments at 1,167 °C were of 2-h duration; experiments at 1,208, 1,306 and 1,402 °C lasted 1 h. Alkali depletion is sensitive to run duration. High-Ca glasses from melting experiments with gypsum lack the MgO enrichment seen in Haiti yellow glass. This may indicate a dolomitic component in the source sediment, as is observed in Cretaceous evaporites around the Chicxulub structure, Mexico²⁵.



tion of the andesite starting material. The composition of the synthetic high-Ca glasses is very similar to the high-Ca impact glasses from Haiti (Table 1). Slight differences between natural and synthetic yellow glasses can be attributed to our choice of an andesite that is slightly more mafic than the natural black glass.

Glasses produced by desilication reactions between carbonate and andesite liquid or evaporite and andesite liquid are identical, with the exception of their sulphur content, and only the high-Ca glasses produced by fusion in presence of evaporite have the high SO_3 content typical of the Haiti high-Ca glass. The synthetic and high-Ca impact glasses show the same enrichment and depletion pattern with respect to the coexisting high-silica glass, including depletion in alkalis (Fig. 2). The alkali depletion is greatest in 2-h experiments, and is not sensitive to temperature. Below 1,150 °C the glasses crystallize melilite, Al-rich clinopyroxene (fassaite), and plagioclase. At higher temperatures, quench crystals of calcic plagioclase and Al-rich augite ($\text{Ca}_{0.96}\text{Mg}_{0.45}\text{Fe}_{0.5}\text{Si}_{1.6}\text{Al}_{0.4}\text{O}_6$) are present in the synthetic charges, which have compositions similar to rare augitic quench microlites found in the high-Ca natural Haiti glass ($\text{Ca}_{1.0}\text{Mg}_{0.7}\text{Fe}_{0.3}\text{Si}_{1.7}\text{Al}_{0.3}\text{O}_6$). The high-Ca impact glass and synthetic glass composition is close to the thermal minimum in the quaternary system $\text{CaO-SiO}_2\text{-Al}_2\text{O}_3\text{-MgO}$ at 1 atm, where clinopyroxene, plagioclase and melilite coexist with liquid at 1,235 °C (refs 14, 15). The high-Ca glass is thus virtually identical to the industrial glass Eutal or E-glass, widely used for glass fibre¹⁶.

The sulphur content of the synthetic high-Ca glasses varies inversely with temperature but is not sensitive to the duration of the experiment. This temperature dependence follows the equation $T\text{ (}^\circ\text{C)} = 1,275.6 + (-245.72(\log \text{ wt\% SO}_3))$, with a correlation coefficient $R^2 = 0.94$. On the basis of these results, the sulphur content of the natural high-Ca glasses from Haiti (0.83 wt% SO_3) indicates a temperature of ~1,300 °C. This inferred temperature and the chemical heterogeneity in Haiti impact glasses have a bearing on their mode of formation. Earlier assumptions of extensive vaporization of target rock in the ejecta plume are not supported by recent experiments, which show that during decompression following a high-velocity impact, the shocked crustal material is mainly liquid¹⁷. The chemical evidence from Haiti glasses is consistent with this. A melt condensed by homogeneous nucleation from rock vapour phase at high temperature should have a very homogeneous composition, because of the high diffusion rates in vapour. The coexistence of two types of glasses in single Haiti impact spherules suggests that they did not form by vapour condensation, but represent ejected melts. The temperature of ~1,300 °C for Haiti impact glasses is also consistent with Raman spectra

of tektites in general, which are similar to spectra of near-liquidus glasses, and thus show no evidence of great superheating¹⁸.

The geological terrane that produced the Haiti impact glasses is probably continental crust of dominantly andesitic composition, overlain by evaporite-rich sediment, which generated the calcic yellow glasses. These results thus indicate an impact site in a shallow sedimentary evaporite basin on a continental shelf. The Manson impact structure in Iowa has a radiometric age that coincides with the K/T boundary¹⁹. The absence of evaporite deposits in the Manson region make this an unlikely source of the Haiti glasses, and comparison of the composition of glass types supports this opinion. Some Manson melts, which have been attributed to a granite origin²⁰, are greatly enriched in REE compared with Haiti glasses and are clearly unrelated. The more common Manson impactites have REE patterns similar to Haiti glasses, but their major and trace element content differs from the black Haiti glass, notably in Na/K ratio (1.8, compared with 0.2–0.6 in Manson rocks), U, Hf, Ta, Sb, Ni and Cr. We point out, however, that these differences are not conclusive, as glasses found within some impact craters can be compositionally distinct from associated tektites²¹.

The Chicxulub structure in Mexico²² provides an excellent match for the geological setting of the K/T boundary impact site predicted from the geochemistry of Haiti impact glasses. Surface features indicate a circular structure, 180 km in diameter, at the coastline of northern Yucatan, beneath thick Tertiary sediments²³. The Yucatan peninsula and Campeche platform was a vast region of shallow-water carbonate and evaporite deposition during the Middle and Late Cretaceous^{24,25}. Between 1 and 3 km of anhydrite and gypsum evaporite formation was deposited, interbedded with dolomitic limestones. Twelve exploration wells reveal thick anhydrite deposits under the Yucatan, as well as andesitic rocks and 'volcanic glass'²⁶. The basement of the Yucatan peninsula is at a depth of 3.2 km and consists of Palaeozoic schists and volcanic rocks with a Rb-Sr age of 410 Myr, which may correspond to the Palaeozoic basement in the Maya mountains to the south²⁷. These basement rocks are a possible source of the high-silica impact glasses in Haiti. Within the Chicxulub structure, the Tertiary sediments overlying the K/T boundary occupy a bowl-shaped depression, ~1 km in depth²⁵. The evaporite formation pinches out or is greatly reduced in thickness within the structure, and the basement rocks form a 1-km-high uplifted central peak. Above the basement, exploration wells recovered 'andesites and tuffs'²⁵ which may be up to 500 m in thickness. Samples from one well contain shocked quartz and microcrystalline andesitic material which may be devitrified or slowly cooled impact melt²⁸.

The possible evaporite nature of the source terrane of the

yellow glass has important implications for atmospheric environmental effects of the impact. Because of the highly volatile behaviour of sulphur at low pressure and high temperature, the observed high sulphur content in the yellow glass is likely to represent only a fraction of the original concentration. Studies of glasses from explosive volcanic eruptions²⁹ show that the fraction of sulphur lost during degassing at the Earth's surface ranges from 0.5 to 0.95. It has been estimated¹ that the mass of yellow glass originally present in the impact layer globally was 5×10^{13} kg. If this represents an evaporite source, then the mass of sulphur degassed into the atmosphere would amount to

$\sim 6.4 \times 10^{15}$ g. This is about two orders of magnitude higher than the sulphur mass emitted from Tambora volcano in 1815, the largest recorded explosive volcanic eruption on Earth. Both theoretical models³⁰ and climate data indicate that large emissions of sulphur dioxide to the stratosphere during volcanic eruptions cause considerable short-term cooling at the Earth's surface²⁹. Extrapolation of the relationship between emitted sulphur mass and the climate effect observed after volcanic eruptions²⁹ indicates that the sulphur mass emitted by impact on an evaporite terrane could result in a hemispheric temperature decrease of the troposphere of $\sim 4^\circ\text{C}$. □

Received 25 April; accepted 17 September 1991.

1. Sigurdsson, H. *et al.* *Nature* **349**, 482–487 (1991).
2. Izett, G. A. *Lunar planet. Sci.* **22**, 625–626 (1991).
3. Oskarsson, N. *et al.* *Lunar planet. Sci.* **22**, 1009 (1991).
4. Taylor, S. R. & McLennan, S. M. *The Continental Crust: its Composition and Evolution* 312 (Blackwell, Oxford, 1985).
5. White, W. M. & Patchett, J. *Earth planet. Sci. Lett.* **67**, 167–185 (1984).
6. Migdisov, A. A., Ronov, A. B. & Grinenko, V. A. in *The Global Geochemical Sulphur Cycle* (eds. Ivanov, M. V. & Freney, J. R.) 25–127 (1983).
7. Clapp, G. E., Holser, W. T., Kaplan, I. R., Sakai, H. & Zak, T. *Chem. Geol.* **28**, 199–260 (1980).
8. Rye, R. O., Luhr, J. F. & Wasserman, M. D. *J. Volcanol. geotherm. Res.* **23**, 109–123 (1984).
9. Swihart, G. H., Moore, P. B. & Callis, E. L. *Geochim. cosmochim. Acta* **50**, 1297–1301 (1986).
10. Wang, Y. L., Liu, Y. G. & Schmitt, R. A. *Geochim. cosmochim. Acta* **50**, 1337–1355 (1986).
11. Palmer, M. R. & Elderfield, H. *Nature* **314**, 526–528 (1985).
12. Butler, G. P. in *The Persian Gulf* (ed. Purser, B. H.), 423–452 (Springer, Berlin, 1973).
13. Cantagrel, J. M., Gourgaud, A. & Robin, C. *Bull. Volcanol.* **47**, 735–748 (1984).
14. Prince, A. T. *J. Am. Ceram. Soc.* **37**, 406 (1954).
15. Osborn, E. F., DeVries, R. C., Gee, K. H. & Kraner, H. M. *Trans. AIME* **200**, 38–39 (1954).
16. Hlavac, J. *The Technology of Glass and Ceramics* 431 (Elsevier, Amsterdam, 1983).
17. Melosh, H. J. & Vickery, A. M. *Nature* **350**, 494–497 (1991).
18. Jakes, P., Sen, S. & Matsuishi, K. *Abstr. Lunar planet. Soc. Conf.* **22**, 633–634 (1991).
19. Hartung, J. B., Kunk, M. J. & Anderson, R. R. *Geol. Soc. Am. Spec. Paper* **247**, 207–221 (1990).
20. Koeberl, C. & Hartung, J. B. *Lunar planet. Sci.* **22**, 733–734 (1991).
21. Delano, J. W. & Lindsley, D. H. *Geochim. cosmochim. Acta* **46**, 2447–2452 (1982).
22. Penfield, G. T. & Camargo, Z. A. *Ann. Meeting Soc. Explor. Geophys. Abstr.* **51**, 37 (1981).
23. Pope, K. O., Ocampo, A. C. & Duller, C. E. *Nature*, **351**, 105 (1991).
24. Bryant, W. R. *et al.* *Am. Assoc. Petrol. Geol. Bull.* **53**, 2506–2542 (1969).
25. Lopez Ramos, E. *Geologia de Mexico* 2, 2nd edn (University of Mexico, Mexico City, 1979).
26. Wilhelm, O. & Ewing, M. *Geol. Soc. Am. Bull.* **83**, 575–600 (1972).
27. Bateson, J. H. & Hall, I. H. S. *Overseas Memoir* 3, *Inst. Geol. Sci.*, 44 (1977).
28. Kring, D. A., Hildebrand, A. R. & Boynton, W. V. *Abstr. Lunar planet. Sci. Conf.* **22**, 755–756 (1991).
29. Sigurdsson, H. *Paleogeogr. Paleoclim. Paleoecon.* **89**, 277–289 (1990).
30. Pollack, J. B. *et al.* *J. geophys. Res.* **81**, 1071–1083 (1976).
31. Mosbah, M., Metrich, N. & Massiot, P. *Nucl. Instrum. Meth.* (in the press).
32. Turpin, L. *et al.* *Contrib. Miner. Petrol.* **104**, 163–172 (1990).
33. Robin, C. *Andesites* (ed. Thorpe, R. S.) 137–147 (Wiley, New York, 1982).

ACKNOWLEDGEMENTS. We thank M. Mosbah for fluorine analysis and P. Allard, M. Arthur and M. Bender for comments.

Disequilibrium carbon and oxygen isotope variations in natural calcite

J. A. D. Dickson

Department of Earth Sciences, University of Cambridge, Downing Street, Cambridge CB2 3EQ, UK

OXYGEN isotope fractionation in the calcite–water system is widely used to investigate the temperature, fluid composition and reaction rate of calcite precipitation, under the assumption that equilibrium is maintained during precipitation. Here I present analyses of coarsely crystalline, inorganic calcite displaying combined growth and sector zoning, which show this assumption to be untrue. Variations in the isotope composition of successive growth zones can be interpreted as due to changing conditions of precipitation, with equilibrium being maintained throughout. But variations across different sectors of the same synchronous growth surface ($\sim 2\%$ for $\delta^{13}\text{C}$ and $\sim 0.9\%$ for $\delta^{18}\text{O}$) record disequilibrium in the same crystal. This type of disequilibrium may be common, as calcite precipitates in a great variety of crystal form combinations; moreover, many noncarbonate minerals also grow as crystals with two or more crystallographic forms. Indeed, Boyd *et al.*¹ reported fractionation of nitrogen isotopes between cubic and octahedral sectors of a synthetic diamond, but I believe that crystallographically controlled isotope fractionation of stoichiometric ions has not hitherto been recognized.

The oxygen isotope fractionation between calcite and water has given rise to the classic palaeotemperature scale. The magnitude of fractionation has been calculated from lattice parameters^{1,2}, determined from analyses of natural skeletal carbonates³ and from inorganically precipitated, synthetic calcium carbonate^{4,5}. The similarity of results from these different approaches indicate isotope equilibrium was achieved. It is known, however, that metabolic processes involved in the pro-

duction of many skeletal carbonates⁷ cause disequilibrium precipitation relative to the aqueous environment that supports the growth of the organisms; rapid crystallization of inorganic carbonate⁸ also causes disequilibrium precipitation. Slow precipitation of inorganic calcite is usually assumed to be an equilibrium process and was used in establishing calcite–water fractionation factors^{1,6}.

Oxygen isotope data from ancient inorganic calcite are usually interpreted either by using the calcite–water fractionation equation to indicate temperature and/or the $\delta^{18}\text{O}$ value of precipitating waters⁹, or by using this equation coupled with a mass-balance equation to assess water/rock reaction^{10,11}. The assumption in these approaches is that there was isotopic equilibrium between carbonate and water during precipitation. This is often thought to be justified if the material is inorganic in origin and coarsely crystalline, indicating slow precipitation.

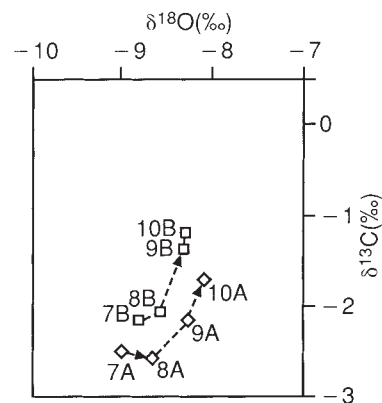


FIG. 1 $\delta^{13}\text{C}$ (‰) plotted against $\delta^{18}\text{O}$ (‰) for Abercrombie calcite. Zones 7 to 10 paired into A and B series from adjacent sectors. $\delta^{13}\text{C}$ values for different crystal faces vary consistently by 0.5‰, whereas $\delta^{18}\text{O}$ values are similar.

TABLE 1 $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ data relative to PDB

Sample no.	$\delta^{13}\text{C}$ (‰)		$\delta^{18}\text{O}$ (‰)			
Abercriban						
7A	-2.5		-9.0			
7B		-2.2		-8.8		
8A	-2.6		-8.7			
8B		-2.1		-8.6		
9A	-2.2		-8.3			
9B		-1.4		-8.3		
10A	-1.7		-8.1			
10B		-1.2		-8.3		
Mean (4)	-2.3	-1.7	-8.5	-8.6		
Bonsall Moor calcite						
	$\delta^{13}\text{C}$ (‰)		$\delta^{18}\text{O}$ (‰)			
	{2131}	{2.9.11.5}	{0111}	{2131}	{2.9.11.5}	{0111}
A01	-4.5			-8.1		
A02		-2.0			-8.4	
A03		-2.3			-8.4	
A04	-4.0			-8.1		
A05	-4.4			-8.0		
A06		-1.7			-8.6	
A07		-2.0			-8.4	
A08	-3.4			-8.0		
A09	-2.2			-9.3		
A10		-1.2			-9.3	
A11		-1.6			-9.2	
A12	-3.3			-8.4		
B13	-3.8			-8.7		
B14		1.0			-9.4	
B15		-1.2			-9.3	
B16	3.7			-7.8		
B17	-2.7			-7.0		
B18		-1.1			-9.2	
B19		-1.4			-9.2	
B20	-3.1			-7.9		
B21	-3.1			-8.7		
B22		-1.1			-9.4	
B23		-1.1			-9.5	
B24	-5.3			-8.7		
B25			-1.4			-9.5
B26			-1.0			-9.3
B27			-2.1			-9.2
C28	-4.8			-8.1		
C29		-2.7			-9.2	
C30	-3.8			-7.5		
C31	-4.6			-7.9		
C32		-2.8			-8.7	
C33	-4.7			-8.0		
C34	-3.7			-7.4		
C35		-3.2			-9.0	
C36	-3.6			-7.4		
C37	-4.3			-8.5		
C38			-1.6			-8.6
C39			-1.9			-8.4
C40			-1.6			-8.8
mean	-3.8	-1.75	-1.59	-8.07	-9.00	-8.98
	(19)	(15)	(6)	(19)	(15)	(6)

Analysis by VG Isogas SIRA Series II mass spectrometer with VG isocarb automated common acid bath system. Calcite samples (~500 mg) baked under vacuum at 420 °C for 30 min to remove organics. Results reported as deviations from PDB standard (‰). Analytical precision (σ) 0.07‰.

Coarse (millimetre-sized) calcite cement crystals from carboniferous limestone, Abercriban (grid ref. SO 065127), South Wales, show combined concentric growth and sector zoning which is revealed by cathodoluminescence or staining¹². Samples for isotope analysis were dissected from thin sections of these crystals and their position recorded relative to zonal patterns shown by staining. Concentric growth zones show isotope variations which could relate to short-term variations in temperature and water composition at isotope equilibrium. But where the crystals show sector zoning (Table 1, samples 7–10, and Fig. 1), $\delta^{13}\text{C}$ for the same growth zone in adjacent sectors varies by ~0.5‰. This implies isotope disequilibrium existed between

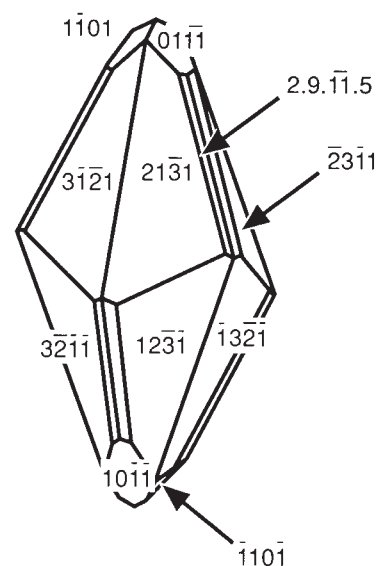


FIG. 2 Idealized drawing of Bonsall Moor calcite showing combination of two scalenohedra, {2131} and {2.9.11.5} with rhombohedral termination {0111}.

different crystal faces at the moment of precipitation. The concentric growth zones have been investigated by laser microprobe which has good spatial resolution¹². Variations of 3‰ for $\delta^{13}\text{C}$ and 6‰ for $\delta^{18}\text{O}$ were found across successive zones only 60 μm apart. Positional sampling errors in such material are clearly important, but they are unlikely to explain the Abercriban data as a consistent $\delta^{13}\text{C}$ shift occurs between adjacent sectors in four successive growth zones.

To investigate whether or not this discovery might have general application, I analysed three calcite crystals (A, B and C) from hydrothermal veins in carboniferous limestone, Bonsall Moor, Derbyshire UK (grid ref. SK 250694). These crystals, 3–7 cm in size, are bounded by crystal faces belonging to three crystallographic forms (Fig. 2). Identification of the poorly developed scalenohedron {2.9.11.5} is approximate, for its faces are curved and angular measurement is therefore inaccurate. Calcite powder for analysis was scraped from the external surface of each crystal face with a steel needle, except for crystal B where material from adjacent paired {2.9.11.5} faces was pooled because of their small size. Because only the crystal surface was sampled, the samples all belong to the same growth zone (the last one). A complete calcite crystal (Fig. 2) will possess 12 faces for each scalenohedron and 6 faces for each rhombohedron developed. Data for individual faces are listed in Table 1, grouped under the general form index to which those faces belong. For instance, samples A01, A04, A05, A08, A09 and A12 come from individual faces {2131}, {2311}, {3211}, {1231}, {1321} and {3121}, which are upper faces of the dominant scalenohedron (Fig. 2) with general form index {2131}. Some sampling error due to uneven depth of scraping below the surface is possible. Except for sample A09 (Table 1), the {2131} samples are differentiated in both $\delta^{13}\text{C}$ (<2.0‰) and $\delta^{18}\text{O}$ (~0.9‰) from the {2.9.11.5} and {0111} samples (Fig. 3), showing significant sectorial fractionation of both isotopes.

The atomic environment for attachment of ions on a growing crystal face will depend on the face's orientation relative to the lattice and the fluid. Cation partitioning on sector calcite has been ascribed to lattice geometry, to the protosite model¹³ and to crystal/ligand field stabilization energies¹⁴. Similar mechanisms are probably responsible for the isotope discrimination between different crystal surfaces.

The effect of face-determined fractionation is not apparent in the experimental work carried out to determine the oxygen

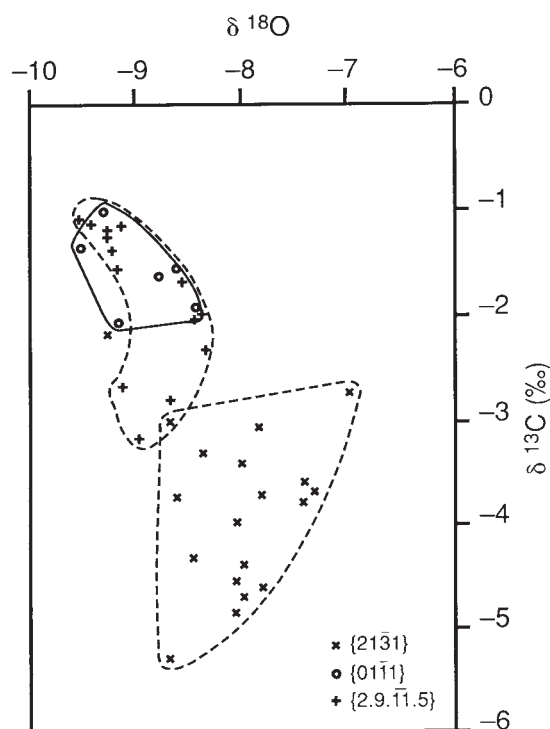


FIG. 3 $\delta^{13}\text{C}$ plotted against $\delta^{18}\text{O}$ for three crystals of Bonsall Moor calcite. Data from three crystallographic forms contained in envelopes, except for sample A09, which is the $\{2131\}$ form.

fractionation in the calcite-water system. One may speculate that the synthetically produced crystals developed either as a single crystallographic form (commonly $\{10\bar{1}1\}$ rhombohedra whose face-dependent fractionation is unknown) with little or no fractionation against equilibrium, or as polyhedral crystals where the effect of different face-dependent fractionation factors cancel one another out, hence accounting for the correspondence between the experimental and theoretically determined fractionation factor for oxygen in calcite. Face-dependent isotope fractionation is clearly important in sampling strategy for stable isotope analysis and raises doubts about the interpretation of existing results that are crystallographically unconstrained. If calcite develops as a single crystallographic form, bulk sampling is satisfactory, but the fractionation of that form against equilibrium must be known before a full interpretation can be made. If calcite crystallizes as a combination of two or more forms, then either intracrystalline sampling is required or the relative proportions of different sectors must be quantified.

The discovery of significant intracrystalline sectorial fractionation in calcite has implications for isotope systems in other crystalline solids that exhibit two or more different crystallographic forms. Boyd *et al.*¹ identified sectorial ($\{100\}/\{111\}$) $\delta^{15}\text{N}$ fractionation of this trace component in synthetic diamond accompanied by a concentration difference ($\{100\} \sim 40$ p.p.m.; $\{111\} \sim 115$ p.p.m.) but were unable to verify carbon fractionation. The face-dependent $\delta^{13}\text{C}$ and $\delta^{18}\text{O}$ fractionation reported here involves stoichiometric components of calcite, the most extensively analysed of all natural minerals. □

10. Prezbindowski, D. R. *Carbonate Cements Spec. Publ.* **36** (eds Schneiderman, N. & Harris, P. M.) 241–264 (SEPM, Tulsa, 1985).
11. Lawrence, J. R. *et al. Init. Rep. Deep-Sea Drilling Proj.* 507–512 (US Government Printing Office, Washington DC, 1976).
12. Banner, J. L. & Hanson, G. N. *Geochim. cosmochim. Acta* **54**, 3123–3137 (1990).
13. Dickson, J. A. D. *et al. Geology* **18**, 809–811 (1990).
14. Reeder, R. J. & Grams, J. C. *Geochim. cosmochim. Acta* **51**, 187–194 (1987).
15. Searl, A. *Mineralog. Mag.* **54**, 501–507 (1990).

ACKNOWLEDGEMENTS. The manuscript was improved by the critical comments of C. P. Smalley and A. Saller.

Early-onset Alzheimer's disease caused by mutations at codon 717 of the β -amyloid precursor protein gene

Marie-Christine Chartier-Harlin, Fiona Crawford, Henry Houlden, Andrew Warren*, David Hughes, Liana Fidani, Alison Goate, Martin Rossor, Penelope Roques, John Hardy & Mike Mullan†

Alzheimer's Disease Research Group, Departments of Biochemistry and Neurology, St Mary's Hospital Medical School, Imperial College, London W2 1PG, UK

* Departments of Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, Baltimore, Maryland 21205, USA

† To whom correspondence should be addressed

A MUTATION at codon 717 of the β -amyloid precursor protein gene has been found to cosegregate with familial Alzheimer's disease in a single family¹. This mutation has been reported in a further five out of ~ 100 families multiply affected by Alzheimer's disease^{1–4}. We have identified another family, F19, in which we have detected linkage between the β -amyloid precursor protein gene and Alzheimer's disease. Direct sequencing of exon 17 (ref. 5) in affected individuals from this family has revealed a base change producing a Val $\rightarrow$ Gly substitution, also at codon 717. The occurrence of a second allelic variant at codon 717 linked to the Alzheimer's phenotype supports the hypothesis that they are pathogenic mutations.

We examined the segregation of familial Alzheimer's disease (FAD) with a highly polymorphic dinucleotide repeat marker GT12 (*D21S210*) which is located close to the β -amyloid precursor protein (APP) gene. In pedigree F19 (Fig. 1), in which FAD was confirmed by autopsy (onset age, 59 ± 4 years), an allele of *D21S210* cosegregates with the disease. Linkage analysis gave a peak lod score between the marker and the disease of 3.02 at a recombination fraction, θ , of zero (Table 1).

We therefore sequenced the APP gene in this family starting with exon 17. Direct sequencing (Fig. 2) revealed a T $\rightarrow$ G transversion at base pair 2,150 changing valine to glycine at codon 717 of the APP transcript 770. Direct sequencing and single-strand conformation analysis (SSCA)^{6,7} show that the mutation cosegregates with the linked marker GT12 and therefore the disease (lod score, 3.66; $\theta = 0$; three affected and 10 unaffected family members). It does not occur in any of the other 24 pedigrees with FAD that we have sequenced, nor does it occur

TABLE 1 Linkage analysis between GT12 and Alzheimer's disease

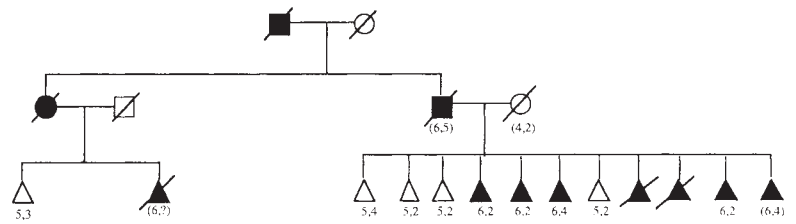
Recombination fraction	0	0.01	0.05	0.1	0.2	0.3	0.4
Lod score	3.02	2.97	2.75	2.47	1.86	1.22	0.6

Lod scores were calculated with seven liability classes modelling age-dependent penetrances from 0.01 to 0.95 with a phenocopy rate of 0.001 and a gene frequency of 0.001 using MLINK from the LINKAGE package¹⁷. (GT12 allele frequencies available from A.W.)

Received 28 May; accepted 16 September 1991.

1. Boyd, S. R., Pillinger, C. T., Milledge, H. J., Mendelsohn, M. J. & Seal, M. *Nature* **331**, 604–607 (1988).
2. McCrea, J. M. *J. chem. Phys.* **18**, 849–857 (1950).
3. Bottinga, Y. *J. chem. Phys.* **72**, 800–808 (1968).
4. Epstein, S. *et al. Geol. Soc. Am. Bull.* **64**, 1315–1326 (1953).
5. Clayton, R. N. *J. chem. Phys.* **34**, 724–726 (1961).
6. Tarutani, T. *et al. Geochim. cosmochim. Acta* **33**, 987–996 (1969).
7. O'Neill, J. R. *et al. J. chem. Phys.* **51**, 5547–5558 (1969).
8. Land, L. S. *Handbook of Environmental Isotope Geochemistry* **3** (eds Fritz, P. & Fontes, J. Ch.) 191–217 (Elsevier, Amsterdam, 1989).
9. Turner, J. V. *Geochim. cosmochim. Acta* **46**, 1183–1191 (1982).

FIG. 1 Diagram of the pedigree together with the GT12 data. Filled symbols represent affected individuals. Numbers designate six alleles of this marker segregating in the pedigree. The pedigree is disguised for reasons of confidentiality. Brackets indicate data derived from reconstruction.



in 52 normal individuals. The occurrence of this mutation is remarkable because it is a change at the same codon as that previously reported¹. Direct sequencing or SSCA of other exons has not revealed any other changes in the APP gene in this family (Fig. 2 legend), but it is known that SSCA does not detect all single base-pair changes⁷. It remains a possibility that other sequence differences in the APP gene exist between affected and unaffected individuals either in exons not yet examined or sequenced or in the promoter.

However, the occurrence of this mutation in a family with histologically confirmed FAD, together with the multiple occurrence of the APP717 Val→Ile mutation, strongly support the hypothesis that these mutations are pathogenic¹. The main features of FAD in F19 and families with the APP717 Val→Ile mutation clinically fulfill NINCDS⁸ criteria for probable Alzheimer's disease and neuropathologically fulfill NINCDS criteria for definite Alzheimer's disease. Clinico-pathological features discriminating between these allelic variants have so far not been detected (M.M., unpublished data). The fact that both mutations are at the same codon may be a coincidence, but this is unlikely. There are two possibilities that may explain our data. First, substitution of the valine residue with Ile or Gly at codon 717 results in increased β -amyloid deposition because of changes in APP processing. Second, variation in the sequence around this position may result in increased translation of APP messenger RNAs, causing Alzheimer's disease by a route analogous to that believed to occur in Down's syndrome^{9,10}.

The mutation described here, like the one reported previously¹, could destabilize a putative stem-loop structure and destroy a possible iron-responsive element between base pairs 2,131 and 2,156⁹. There are several other possible mutations that could also disrupt this structure, many of which would be silent at the protein level. Both the Val→Ile and the Val→Gly mutations involve changes in amino acid 717; no silent changes or changes to other amino acids have yet been reported. Examination of sequence data from 10 other mammalian species¹¹ shows that although the valine residue at codon 717 is conserved in all species, the putative stem-loop structure

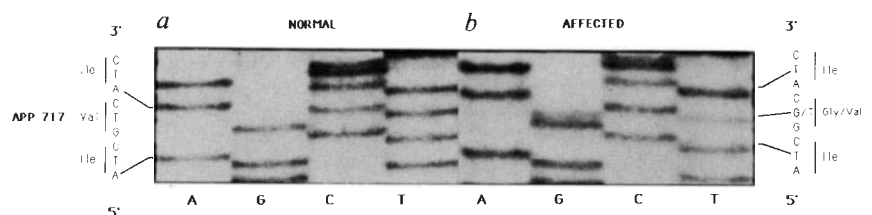
postulated from the human sequence⁹ would not be predicted to occur in either cattle or sheep; and in pig and mouse the consensus sequence for iron-responsive elements¹² is not present. Finally, such stem-loop structures are believed to modulate gene translation by altering mRNA stability¹². However, *in situ* hybridization suggests that amounts of APP mRNAs are not grossly altered in the brain of an individual with APP717 Val→Ile¹³. For these reasons, we suggest that alterations in the rate of APP translation caused by these mutations are not likely to be the key to their pathogenicity. The fact that the mutations can convert Val to either Gly or Ile suggests that neither side-chain hydrophobicity nor bulk are responsible for the pathogenicity of these mutations.

The principal fate of APP involves cleavage within the β -amyloid fragment¹⁴. However, deposition of β -amyloid occurs to a small extent under normal conditions in the absence of any mutations. Therefore an alternative pathway of APP processing must occur that involves cleavage on either side of the β -amyloid region of APP. Amyloid molecules isolated from the brain of a patient with Alzheimer's disease show N- and C-terminal heterogeneity^{15,16}, suggesting that this alternative pathway involves either sequence-specific proteolysis followed by exopeptidase activity (creating end-heterogeneity), or that it is not sequence-specific. We suggest that proteolysis on either side of the β -amyloid region of APP is enhanced by these mutations, increasing the rate of β -amyloid deposition. Such mutations might increase β -amyloid formation by providing a poorer substrate for the main intra-amyloid proteolytic pathway or a better substrate for the alternative pathway. The identification and characterization of further mutations in familial Alzheimer's disease, together with work aimed at elucidating APP processing, should help to resolve the mechanisms by which these mutations are pathogenic and may suggest other causes of β -amyloid deposition¹⁶.

Note added in proof: A third mutation at codon 717 of the APP770 transcript associated with Alzheimer's disease has been reported. The predicted amino-acid change is valine to phenylalanine²⁰. □

FIG. 2 APP exon-17 sequences in an affected and an unaffected member of F19. In the affected member, there is a T→G transition at nucleotide position 2,150.

METHODS. Amplification of exon 17 was as described¹⁸, with the following modifications: (1) the amplification primer sequences were 5'-ATAACCTCATCCAAATGTCCTC-3' and 5'-GTAACCCAAGCATCATGGAAG-3'; (2) PCR conditions were 94°C for 10 min, then 35 cycles of 60°C for 1 min; 72°C, 1 min; 94°C, 1 min, followed by 72°C for 5 min. The second primer (50 pmol) was used to generate single-stranded product which was then purified and sequenced with the Sequenase kit (2.0) (USB) using a primer of sequence 5'-AAATGAAATTCT-TCTAATTGCG-3'. The presence of the T→G transversion creates artefacts on gels which are resolved by inclusion of inosine (Sequenase kit) in the



sequencing reaction. Direct sequencing of exons 7 and 16 from affected individuals from F19 showed that these were of normal sequence and SSCA failed to identify changes in exons 2, 3, 7, 9, 12, 13, 15 or 18. SSCA of exon 17 detects APP693 (ref. 19), and both APP717 Val→Ile and Val→Gly (conditions available from the authors).

Received 16 August; accepted 4 October 1991.

1. Goate, A. *et al. Nature* **349**, 704–706 (1991).
2. Hardy, J. *et al. Lancet* **337**, 1342–1343 (1991).
3. van Duijn, C. M. *et al. Lancet* **337**, 978 (1991).
4. Naruse, S. *et al. Lancet* **337**, 978–979 (1991).
5. Yoshikai, S., Sasaki, H., Dohura, K., Furuya, H. & Sakaki, Y. *Gene* **87**, 257–263 (1990).
6. Orita, M., Suzuki, Y., Sekiya, T. & Hayashi, K. *Genomics* **5**, 874–879 (1989).
7. Orita, M., Sekiya, T. & Hayashi, K. *Genomics* **8**, 271–278 (1990).
8. McKhann, G. *et al. Neurology* **34**, 939–944 (1984).
9. Tanzi, R. E. & Hyman, B. T. *Nature* **350**, 564 (1991).
10. Rumble, B. *et al. New Engl. J. Med.* **320**, 1446–1452 (1989).
11. Johnstone, E. M., Chaney, M. O., Norris, F. H., Pascual, R. & Little, S. P. *Molec. Brain Res.* **10**, 299–305 (1991).
12. Klausner, R. D. & Harford, J. B. *Science* **246**, 870–872 (1989).
13. Harrison, P. J., Barton, A. J. L. & Pearson, R. C. A. *Neuro Rep.* **2**, 152–154 (1991).

14. Esch, F. S. *et al. Science* **248**, 1122–1124 (1990).
15. Master, C. *et al. Proc. natn. Acad. Sci. U.S.A.* **82**, 4245–4249 (1985).
16. Glenner, G. G. & Wong, C. W. *Biochem. biophys. Res. Commun.* **120**, 885–890 (1984).
17. Lathrop, G., Lalouel, J. M., Julier, C. & Ott, J. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3443–3446 (1984).
18. Chartier Harlin, M. C. *et al. Neurosci. Lett.* **129**, 134–135 (1991).
19. Levy, E. *et al. Science* **248**, 1124–1126 (1990).
20. Murrell, J., Farlow, M., Ghetti, B. & Benson, M. *Science* **253**, 97–98 (1991).

ACKNOWLEDGEMENTS. We thank R. Harris for referring the family to us, D. Bowen and D. Mann for samples of brain tissue from an affected individual and D. Mann for confirmation of diagnosis, S. Antonarakis and M. McInnis for sharing data with us before publication, and M. Palmer, F. Chartier and F. Ashall for discussion. This work was supported by the Mental Health Foundation, Research into Ageing, the Alzheimer's Disease Society, the Royal Society, the IPSEN Foundation, the Wellcome Trust and the Medical Research Council.

Mutations in the channel domain alter desensitization of a neuronal nicotinic receptor

F. Revah*, D. Bertrand†, J.-L. Galzi*,
A. Devillers-Thiéry*, C. Mulle*, N. Hussy†,
S. Bertrand†, M. Ballivet‡ & J.-P. Changeux*

* Neurobiologie Moléculaire, Unité de Recherche Associée au Centre National de la Recherche Scientifique D 1284, Institut Pasteur, 25 rue du Dr Roux, 75724 Paris Cedex 15, France

† Département de Physiologie, Centre Médical Universitaire (Faculté de Médecine), and ‡ Département de Biochimie, Université de Genève, 1211 Genève 4, Switzerland

A VARIETY of ligand-gated ion channels undergo a fast activation process after the rapid application of agonist and also a slower transition towards desensitized or inactivated closed channel states when exposure to agonist is prolonged^{1–5}. Desensitization involves at least two distinct closed states in the acetylcholine receptor, each with an affinity for agonists higher than those of the resting or active conformations^{1–5}. Here we investigate how structural elements could be involved in the desensitization of the acetylcholine-gated ion channel from the chick brain α -bungarotoxin sensitive homo-oligomeric α 7 receptor^{6,7}, using site-directed mutagenesis and expression in *Xenopus* oocytes. Mutations of the highly conserved leucine 247 residue^{8,9} from the uncharged MII segment of α 7 suppress inhibition by the open-channel blocker QX-222 (ref. 10), indicating that this residue, like others from MII (refs 11–22), faces the lumen of the channel. But, unexpectedly, the same mutations decrease the rate of desensitization of the response, increase the apparent affinity for acetylcholine and abolish current rectification. Moreover, unlike wild-type α 7, which has channels with a single conductance level, the leucine-to-threonine mutant has an additional conducting state active at low acetylcholine concentrations. It is possible that mutation of Leu 247 renders conductive one of the high-affinity desensitized states of the receptor.

All mutants were constructed at position Leu 247 (L247S, L247T, L247V, L247F; one-letter amino-acid code) yielded func-

tional receptors (Fig. 1b). There were no significant differences in Hill coefficient or in sensitivity to α -bungarotoxin between α 7 wild type and the mutants (Fig. 1c), indicating that the overall structure of the protein was not significantly altered. The open-channel blocker QX-222 (ref. 10) inhibited the ionic response of the α 7 wild-type acetylcholine receptor (AChR) in a voltage-dependent manner (Fig. 1e). Although the block persists when Leu 247 is substituted by a hydrophobic residue, it is abolished by polar mutations (T, S) (Fig. 1e; Table 1). This indicates that Leu 247 points towards the interior of the channel, and interacts hydrophobically with QX-222.

Mutation of Leu 247 modifies other features of the physiological response of the receptor to acetylcholine (ACh). The magnitude of these changes always increases in the order F, V, T, S, with the polar mutations giving rise to the most pronounced modifications. First, the time course of the permeability response to ACh of the L247 mutants differs from the wild type (Fig. 1b; Table 1). The rates of desensitization decrease with the size and/or hydrophobicity of the amino-acid side chain; furthermore, whereas all wild-type channels quickly close after removal of the agonist, a substantial residual current can be recorded for several seconds with the L247V, L247T and L247S mutants. Second, mutation of Leu 247 increases the sensitivity of the receptor to ACh (Fig. 1c): the nonpolar mutants were 4 to 18 times, and the polar mutants ~160 times, more sensitive to the agonist than the wild type. Third, the strong current rectification observed at positive potentials for the wild type, L247F and L247V receptors is abolished for the polar mutants (Fig. 1d).

The range of ACh concentrations that activate the L247 mutants was compared with that eliciting desensitization of the wild-type receptor. We examined the effect of pre-exposure to low agonist doses on a test response of α 7 wild type; Fig. 2 shows that the concentrations of ACh that caused desensitization of the wild-type receptor without evoking an ionic response activated the L247T mutant.

To analyse further the effect of polar substitutions of Leu 247, we compared the single-channel conductance of the wild-type and L247T receptors. In the cell-attached patch configuration, recordings from oocytes expressing α 7 wild-type receptor reveal a single class of ACh-activated channels (45 pS, $n = 3$; Fig. 3). In contrast, oocytes injected with L247T complementary DNA show two conducting states (Fig. 3). One of these has features

TABLE 1 Data for wild-type and L247 mutant receptors

AChR type	Peak current at 100 μ M ACh (nA)	Per cent of current at 2 s (100 μ M ACh)	Half decay time (s) of response (100 μ M ACh)	EC ₅₀ (μ M)	Per cent of block by 100 μ M QX-222 (–120 mV)
α 7 WT	1,745	7.4	0.6	115	85
α 7 L247F	157	21.8	0.4	30	58
α 7 L247V	2,397	78.4	5.5	6.2	77
α 7 L247T	1,664	77.0	9.5	0.7	0
α 7 L247S	1,810	77.4	15.0	0.8	0

See Fig. 1 for methods. Similar recordings were obtained for every mutation in at least $n = 5$ cells from more than one batch of oocytes ($n > 500$ for wild type (WT); $n > 50$ for L247T).

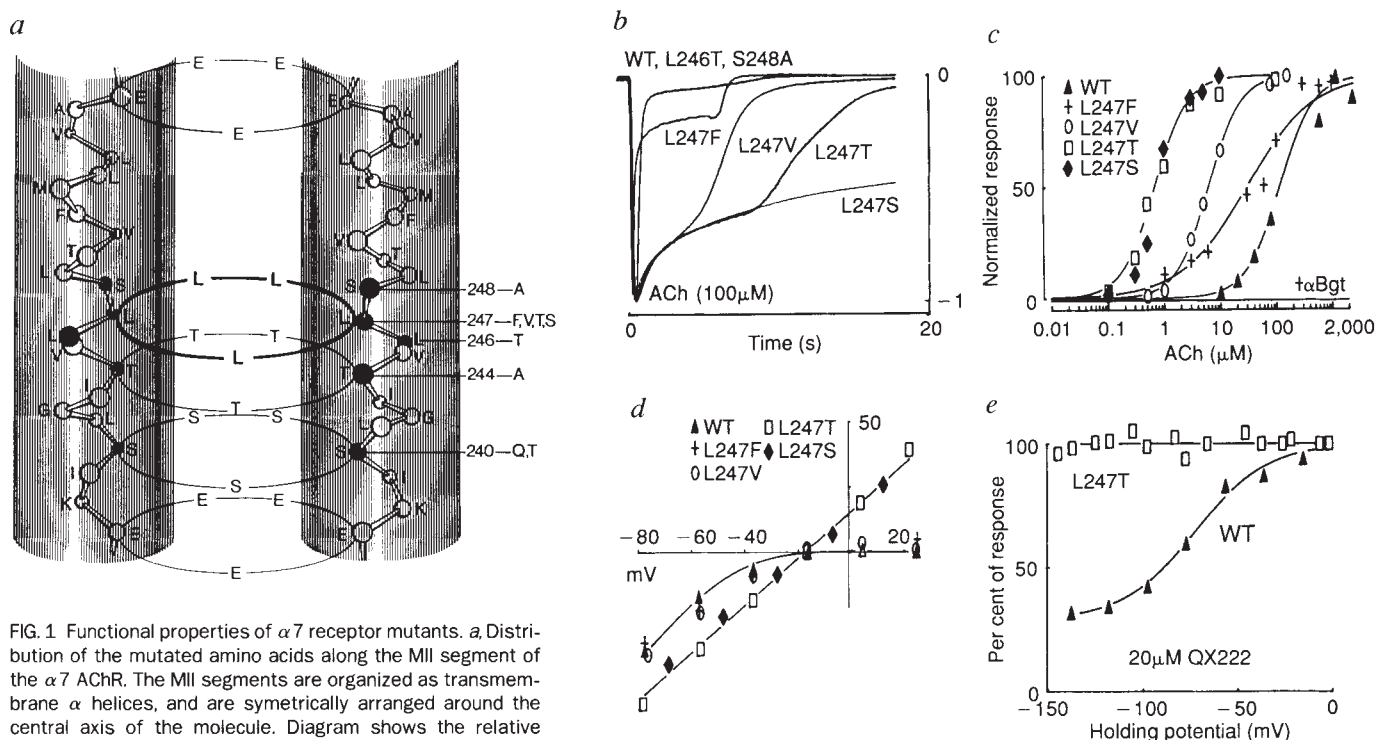


FIG. 1 Functional properties of $\alpha 7$ receptor mutants. **a**, Distribution of the mutated amino acids along the M1 segment of the $\alpha 7$ AChR. The M1 segments are organized as transmembrane α helices, and are symmetrically arranged around the central axis of the molecule. Diagram shows the relative position of the α -carbons of the amino acids (one-letter code). Mutated residues are shown as filled spheres; their position in the $\alpha 7$ sequence and the amino acids with which they have been substituted are indicated on the right. Rings of negatively charged residues framing M1 (positions 237 and 258), and polar (240, 244) or hydrophobic (247) amino acids within M1, which were identified by labelling with non-competitive blockers^{11–16} or previous mutagenesis experiments^{17–22}, are also shown. **b**, Normalized currents evoked by 100 μ M ACh in *Xenopus* oocytes expressing $\alpha 7$ wild type (WT) and $\alpha 7$ L247F, L247V, L247T, L247S, L246T, S248A receptors (where letters represent amino acids at positions given by the numbers) are superimposed. Cells were held at -100 mV and currents were evoked by a 5-s ACh application. Peak currents obtained for the L247 mutants are given in Table 1. The normalized responses to ACh from WT, L246T and S248A were indistinguishable. **c**, ACh dose-response relationships of $\alpha 7$ WT and $\alpha 7$ L247F, L247V, L247T, L247S receptors. Responses evoked by 3-s of increasing ACh concentrations were measured in several cells (5–12) held at -100 mV. Peak currents were then normalized to the maximum values, averaged and plotted as a function of the logarithm of the ACh concentration. Lines are best fits obtained with the empirical Hill equation adjusted by weighted least-square algorithm. Apparent 50% effective concentration (EC_{50}) and Hill coefficients of these receptors are respectively: WT, 115 μ M and 1.4; L247F, 30 μ M and 0.7; L247V, 6.2 μ M and 1.5; L247T, 0.7 μ M and 1.5; L247S, 0.8 μ M and 2.0. The value of the Hill coefficient of the L247F mutant could reflect an altered response due to the extremely fast desensitization of this receptor type. The lower trace

indicates the value of residual WT or L247T currents measured after 30-min exposure to 10 nM α -bungarotoxin. **d**, Current-voltage relationships of $\alpha 7$ WT and $\alpha 7$ L247F, L247V, L247T, L247S receptors. Peak currents evoked by short application of ACh (100 μ M for $\alpha 7$ WT, L247F, L247V, and 0.3 μ M for L247T, L247S) were measured at several holding potentials. Values were then normalized with respect to the peak current obtained at -140 mV and plotted as a function of the holding potential. **e**, Block of ACh-evoked currents by QX-222. Currents evoked by a 3-s pulse of 100 μ M ACh were measured in the absence or presence of 20 μ M QX-222 at several holding potentials. Per cent of response was determined as the ratio of these two ACh-evoked currents and plotted as a function of voltage. For the WT the line is the best fit of the theoretical equation describing voltage-dependent block, y (refs 23, 24), with $y = 100 - 71 / (1 + 44 \exp(x/19))$, where x is the membrane potential.

METHODS. Mutagenesis: The 1,057-base pair (bp) *Pst*I (nucleotide -153)–*Bgl*II (nucleotide 904) and the 2,122-bp *Bgl*II (nucleotide 904)–*Eco*RI (nucleotide 3,026) fragments from plasmid flip $\alpha 7\Delta$ (ref. 7), were ligated and subcloned between the *Pst*I and *Eco*RI sites of bluescript KS+ vector (Stratagene) to permit single-stranded DNA synthesis. The promoter from SV40 virus was introduced in the same vector between *Not*I and *Spe*I sites. Mutants were prepared using an oligonucleotide-directed mutagenesis kit (Amersham) according to the manufacturer's instructions and their sequence systematically checked. For electrophysiology methods, see legend to Fig. 3.

FIG. 2 Desensitization of $\alpha 7$ WT receptor and activation of L247T at low agonist concentrations. **a**, Currents evoked by a 3-s application of 100 μ M ACh recorded in $\alpha 7$ WT receptor with or without a prepulse (8-s) of low ACh concentration (upper panel). Recordings obtained in an oocyte expressing L247T receptor to two ACh concentrations (in the range of that used for the prepulse on the WT receptor), are superimposed in the lower panel. **b**, Dose-response relationship of $\alpha 7$ WT and L247T receptors are compared with $\alpha 7$ WT desensitization. Dose-response curves are the same as those shown in Fig. 1c. Desensitization of $\alpha 7$ WT receptor was determined by measuring the peak amplitude of the currents evoked by a test pulse of 100 μ M ACh after a desensitizing ACh prepulse (protocol as in **a**). Normalized currents were then plotted as a function of prepulse ACh concentration, and fitted with a Hill equation (50% inhibitory concentration, $IC_{50} = 14$ μ M). Recordings performed in several oocytes displayed identical features.

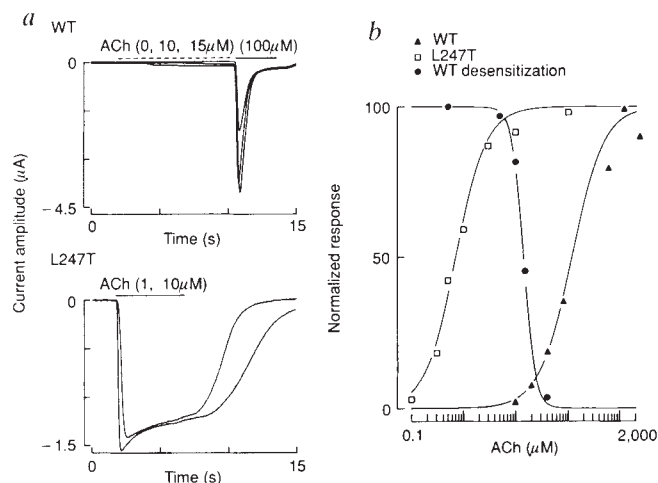


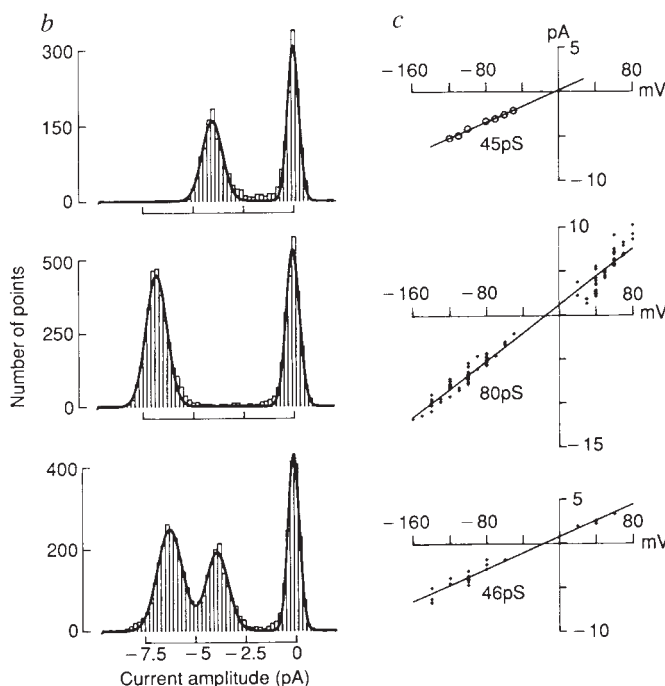
FIG. 3 Single-channel conductances of $\alpha 7$ WT receptors. **a**, Top, typical single-channel recordings from WT and L247T receptors. Recordings of $\alpha 7$ WT receptor were obtained in the cell-attached configuration (10 μ M ACh in the pipette solution; holding potential is -100 mV below rest) and revealed a single population of channels. Middle, single-channel currents of higher amplitude were observed when outside-out patches from oocytes expressing the $\alpha 7$ L247T receptor were challenged with low concentrations of ACh (<3 μ M). Bottom, increasing the ACh concentration above 10 μ M resulted in the appearance of smaller currents (different patch from above). Holding potential of excised patches is -100 mV. Similar channels are observed in cell-attached patches from oocytes expressing mutant receptors. None of the channels are found in the absence of ACh or in non-injected oocytes. **b**, All-point amplitude histograms of the three patches shown in **a**, obtained by cumulative analysis of at least 1.5-s recordings. Gaussian fits give amplitudes of -4.0 pA (WT), -6.8 pA (L247T, 0.1 μ M ACh), -6.3 and -3.9 pA (L247T, 20 μ M ACh). Histograms were binned at intervals of 0.2 pA and best gaussian fits obtained using a weighted least-square algorithm (Simplex). **c**, Top, current-voltage (I - V) relationship of the WT receptor from the same patch shown in **a**. No activity was detected at depolarized potentials. Holding potential is expressed relative to the resting potential of the oocyte. Linear regression gives a slope conductance of 45 pS (correlation coefficient, $R=0.99$) and a reversal potential of -4 mV below rest. Data from all cell-attached patches could not be compiled on the same graph owing to the variability in oocyte resting potential. I - V curves of the large conductance (middle panel; 23 patches) and the small conductance channels (bottom panel; 7 patches) of mutant receptors give slope conductance of 80 pS ($R=0.99$) and 46 pS ($R=0.98$) and reversal potentials of -15 and -17 mV, respectively. Deviation from linearity at highly positive potentials for the 80 pS channel may reflect small outward rectification.

METHODS. Oocytes were prepared, injected and recorded as described²⁵. Recordings were typically made two or three days after injection. Cells were always clamped at -100 mV and challenged by 5-s ACh applications (unless otherwise indicated). Quality and response of oocytes used for single channel recordings were first assessed with the dual electrode voltage clamp. Cells were incubated for 20 min in stripping solution (OR2, 2 $\times$) and the vitelline membrane was removed with fine forceps. Cells were placed in 35-mm Petri

of the wild-type open channel: its conductance was 46 pS ($n=7$) and was detected only at ACh concentrations above 10 μ M. The other had a conductance of 80 pS ($n=23$) and is observed at ACh concentrations as low as 0.1 μ M. As the higher conductance alone is observed at low agonist concentrations (<3 μ M), it cannot represent the activation of two 46 pS channels. These results indicate that mutation of Leu 247 does not significantly alter the conductance of the open state observed in the wild-type receptor, but gives rise to an additional conducting state with higher affinity for ACh.

We also mutated residues adjacent to Leu 247 (L246T, S248A; Fig. 1a, b), or others presumed to lie on the same meridian of the MII helix (S240T, S240Q, T244A; results not shown). None of these substitutions has an effect comparable to Leu 247 mutations on the time course and sensitivity of the response to acetylcholine.

In summary, mutation of Leu 247 in the ion channel modifies the effect of a channel blocker, and concomitantly increases agonist sensitivity, reduces the rates of desensitization, and gives rise to a new elementary conductance. One reason for these effects could be that Leu 247 blocks the channel when the



dishes and superfused with OR2 medium during the entire course of the experiment. Because of apparent clustering of the channels in the oocyte membrane, AChRs were first localized using a puffing electrode and whole-cell recording. WT receptors were more difficult to localize than the mutants probably because of their different properties (such as rapid desensitization). Moreover, single-channel activity of $\alpha 7$ receptor could never be recorded on patches excised from oocytes. Therefore, WT receptors were investigated exclusively in the cell-attached configuration, with a non-desensitizing ACh concentration (10 μ M) in the recording pipette. Even then, only a few patches showed WT receptor activity (5% of the patches), and conductance could be precisely measured in three of them. In contrast, oocytes expressing mutated receptors generally displayed larger macroscopic currents and clusters of receptors were easier to localize. They were therefore characterized in the excised, outside-out configuration that allows a better control of solutions on both sides of the patch as well as of membrane potential. Of the excised patches, 17% had ACh-activated channel activity. Control experiments showed that similar channels were also obtained under the cell-attached mode. Pipette contained OR2 medium in the cell-attached configuration and (in mM) KF, 80; EGTA, 5; MgCl_2 , 0.5; HEPES, 5, pH 7.4, in the excised outside-out mode. ACh was either included in the pipette or the bath solution, or applied through a puffer pipette. Data were filtered at 1.5 kHz and sampled at 5 kHz.

receptor is in one of its high-affinity desensitized states^{1-5,8}, when its replacement by a smaller hydrophilic amino acid would render this high-affinity state permeable. It follows that low concentrations of ACh which desensitize the wild-type receptor should elicit ionic currents when applied to the L247 mutant by activating the new conducting state, as we observed. Our data illustrate the importance of $\alpha 7$ Leu 247 in ion channel function and desensitization mechanisms. Because it is highly conserved in all nicotinic, GABA_A and glycine receptors, this residue may play a similar role in all these ligand-gated ion channels. □

Received 20 May; accepted 28 August 1991.

- Katz, B. & Thesleff, S. *J. Physiol. Lond.* **138**, 63-80 (1957).
- Heidmann, T. & Changeux, J. P. *Eur. J. Biochem.* **94**, 281-296 (1979).
- Sakmann, B., Patlak, J. & Neher, E. *Nature* **286**, 71-73 (1980).
- Neubig, R., Boyd, N. & Cohen, J. B. *Biochemistry* **21**, 3460-3467 (1982).
- Heidmann, T., Bernhardt, J., Neumann, E. & Changeux, J. P. *Biochemistry* **22**, 5452-5459 (1983).
- Schoepfer, R., Conroy, W. G., Whiting, P., Gore, M. & Lindstrom, J. M. *Neuron* **5**, 35-48 (1990).
- Couturier, S. et al. *Neuron* **5**, 847-856 (1990).
- Changeux, J. P. in *Fidia Research Foundation Neuroscience Award Lectures* **4**, 21-168 (Raven, New York, 1990).
- Betz, H. *Neuron* **5**, 383-392 (1990).
- Neher, E. & Steinbach, J. *J. Physiol. Lond.* **277**, 153-176 (1978).

11. Giraudat, J., Dennis, M., Heidmann, T., Chang, J. Y. & Changeux, J. P. *Proc. natn. Acad. Sci. U.S.A.* **83**, 2719–2723 (1986).
12. Giraudat, J. *et al.* *Biochemistry* **26**, 2410–2418 (1987).
13. Giraudat, J. *et al.* *FEBS Lett.* **253**, 190–198 (1989).
14. Revah, F. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **87**, 4675–4679 (1990).
15. Oberthur, W. *et al.* *EMBO J.* **5**, 1815–1819 (1986).
16. Hucho, F., Oberthur, W. & Lottspeich, F. *FEBS Lett.* **205**, 137–142 (1986).
17. Imoto, K. *et al.* *Nature* **324**, 670–674 (1986).
18. Imoto, K. *et al.* *Nature* **335**, 645–648 (1988).
19. Cooper, E., Couturier, S. & Ballivet, M. *Nature* **350**, 235–238 (1991).
20. Leonard, R. J., Labarca, C. G., Charney, P., Davidson, N. & Lester, H. A. *Science* **242**, 1578–1581 (1988).
21. Charney, P. *et al.* *Neuron* **4**, 87–95 (1990).
22. Villarroel, A., Herlitz, S., Koenen, M. & Sakmann, B. *Proc. R. Soc. B* **243**, 69–74 (1991).
23. Woodhull, A. M. *J. gen. Physiol.* **61**, 687–708 (1973).
24. Johnson, J. W. & Ascher, P. *Biophys. J.* **57**, 1085–1090 (1990).
25. Bertrand, D., Cooper, E., Valera, S., Rungger, D. & Ballivet, M. in *Methods in Neuroscience* **4** (ed. Conn, P. M.) 174–193 (Academic, San Diego, 1991).

ACKNOWLEDGEMENTS. We thank P. Ascher for discussion, J. Hill and R. Miles for reading the manuscript, and J. Igolen for synthesis of QX-222. This work was supported by grants from the Association Française contre les Myopathies, the Collège de France, the Centre National de la Recherche Scientifique, the Ministère de la Recherche, the Institut National de la Recherche Médicale, the Direction des Recherches Etudes et Techniques, the Swiss National Science Foundation (D.B. and M.B.) and the Roche Foundation (N.H.).

Inositol trisphosphate-dependent periodic activation of a Ca^{2+} -activated K^+ conductance in glucose-stimulated pancreatic β -cells

Carina Ämmälä, Olof Larsson, Per-Olof Berggren*, Krister Bokvist, Lisa Juntti-Berggren*, Henrik Kindmark* & Patrik Rorsman†

Department of Medical Physics, Gothenburg University, Box 33031, S-400 33 Gothenburg, Sweden

GLUCOSE-STIMULATED insulin secretion is associated with the appearance of electrical activity in the pancreatic β -cell. At intermediate glucose concentrations, β -cell electrical activity follows a characteristic pattern of slow oscillations in membrane potential on which bursts of action potentials are superimposed¹. The electrophysiological background of the bursting pattern remains unestablished. Activation of Ca^{2+} -activated large-conductance K^+ channels (K_{Ca} channel²) has been implicated in this process³ but seems unlikely in view of recent evidence demonstrating that the β -cell electrical activity is unaffected by the specific K_{Ca} channel blocker charybdotoxin⁴. Another hypothesis postulates that the bursting arises as a consequence of two components of Ca^{2+} -current inactivation⁵. Here we show that activation of a novel Ca^{2+} -dependent K^+ current in glucose-stimulated β -cells produces a transient membrane repolarization. This interrupts action potential firing so that action potentials appear in bursts. Spontaneous activity of this current was seen only rarely but could be induced by addition of compounds functionally related to hormones and neurotransmitters present in the intact pancreatic islet. K^+ currents of the same type could be evoked by intracellular application of GTP, the effect of which was mediated by mobilization of Ca^{2+} from inositol 1,4,5-trisphosphate (InsP_3)-sensitive intracellular Ca^{2+} stores. These observations suggest that oscillatory glucose-stimulated electrical activity, which is correlated with pulsatile release of insulin⁶, results from the interaction between the β -cell and intra-islet hormones and neurotransmitters. Our data also provide evidence for a close interplay between ion channels in the

plasma membrane and InsP_3 -induced mobilization of intracellular Ca^{2+} in an excitable cell.

Figure 1a (i) shows a membrane potential recording from a β -cell in a large aggregate stimulated with 10 mM glucose. The cell produces bursts of action potentials which are separated by repolarized intervals. Voltage-clamp recording from the same cell revealed that this type of electrical activity was associated with outward current oscillations. Such electrical activity and membrane currents have been observed in a total of only two cells (<5%). In the vast majority of β -cells, the resting membrane conductance does not fluctuate spontaneously and electrical activity in the presence of 10 mM glucose consists of either continuous electrical activity or very long bursts of action potentials (Fig. 1c, trace i; compare ref. 7). Rapid oscillations in membrane potential could, however, be induced by adding either 3 μM carbamylcholine (Fig. 1b) or 1 mM dibutyryl cyclic AMP (db-cAMP) (Fig. 1c; compare ref. 8), compounds functionally related to hormones and neurotransmitters that are present within the intact pancreatic islet⁹. The oscillations induced by carbamylcholine were paralleled by periodic increases in the cytoplasmic concentration of free Ca^{2+} [Ca^{2+}]_i. In the presence of either db-cAMP or carbamylcholine, transient and repetitive membrane currents were observed (Fig. 1c, trace ii, and Fig. 1d). The actions of db-cAMP and carbamylcholine interact and the frequency of the current oscillations is markedly increased in the presence of both compounds (Fig. 1c, trace ii). The K^+ currents induced by carbamylcholine occur independently of whether extracellular Ca^{2+} is present (Fig. 1d, trace i) or not (trace ii).

Figure 2 shows that K^+ -current oscillations, similar to those observed in β -cells stimulated with glucose and carbamylcholine or db-cAMP, can be evoked by intracellular application of GTP (not shown) or its non-hydrolysable analogue GTP γS (Fig. 2a). Under the latter experimental conditions, repetitive oscillations were maintained for up to 45 min and occurred at a frequency of between one and four per min in different cells. Such currents were also observable in cells dialysed with medium free of exogenous guanine nucleotides but then only immediately after establishment of the whole-cell configuration and limited to a maximum of three oscillations (not shown). The frequency of the oscillations induced by GTP γS can be modulated by addition of carbamylcholine (Fig. 2b). In five different cells, the mean interval between two oscillations in the absence of the agonist was 28 ± 7 s. After addition of 3 μM carbamylcholine, the interval decreased to 14 ± 3 s ($P < 0.05$). Addition of a higher concentration of carbamylcholine ($\geq 150 \mu\text{M}$) first initiated high frequency oscillations but then abolished activity. Carbamylcholine had no effect when added in the absence of intracellular GTP γS . These observations indicate that GTP γS and carbamylcholine induce K^+ -current oscillations by the same mechanism(s). Figure 2d shows that the membrane conductance changes induced by GTP γS are paralleled by variations in [Ca^{2+}]_i. GTP γS -induced [Ca^{2+}]_i transients have previously been reported even when the β -cell is voltage-clamped at voltages too negative for Ca^{2+} -channel activation^{10,11}. Indeed, the K^+ -current oscillations induced by GTP γS do not seem to require Ca^{2+} influx as they occur even after removal of extracellular Ca^{2+} (Fig. 2c). In four different cells, the mean interval between two successive oscillations in the presence and absence of extracellular Ca^{2+} averaged 28 ± 4 and 25 ± 2 s, respectively. The peak amplitude was also unaffected and amounted to 49 ± 7 and 50 ± 5 pA before and after removal of Ca^{2+} . Figure 2e (trace i) shows a membrane potential recording from a β -cell dialysed with GTP γS . The cell oscillates between a repolarized potential and a depolarized plateau potential on which action potentials were superimposed. The variations of the membrane potential are associated with the activation of a rapidly activating and spontaneously inactivating K^+ current (Fig. 2e, trace ii). Extracellular application of db-cAMP during standard whole-cell conditions initiated currents, with properties similar to those

* Permanent address: Department of Endocrinology, Karolinska Institute, Box 60500, S-104 01 Stockholm, Sweden.

† To whom correspondence should be addressed.

induced by GTP γ S, even in the absence of exogenous GTP γ S. These currents were resistant to Ca $^{2+}$ removal, suggesting that elevation of cyclic AMP also leads to mobilization of intracellular Ca $^{2+}$ (data not shown; compare ref. 12).

The pharmacological properties of the GTP γ S-induced currents are shown in Fig. 3. As expected from the fact that they were associated with large [Ca $^{2+}$] $_i$ transients, much of the current recorded at zero mV membrane potential reflects activation of large-conductance K $_{Ca}$ channels, as evidenced by its sensitivity to TEA 13,14 and the selective blocker charybdotoxin 15 (Fig. 3a). At the more physiological membrane potential -40 mV, however, the observed currents were largely unaffected by these pharmacological agents (Fig. 3b). This is in keeping with the

observation that the gating of the K $_{Ca}$ channel in the β -cell shows strong voltage dependence 2 . The currents were also resistant to apamin (1–5 μ M), a blocker of low-conductance Ca $^{2+}$ -activated K $^+$ channels 16 (not shown). These findings suggest that the current observed at the negative membrane potentials reflects activation of a novel Ca $^{2+}$ -activated K $^+$ conductance, which flows through K $^+$ channels distinct from the large-conductance K $_{Ca}$ channels previously characterized in the β -cell 2 . The observed pharmacological properties are those expected from membrane potential recordings of a current involved in the generation of the oscillatory electrical activity 15,17 . Addition of tolbutamide (500 μ M; an inhibitor of the K $_{ATP}$ channel 18) had little effect on the GTP γ S-induced current but suppressed

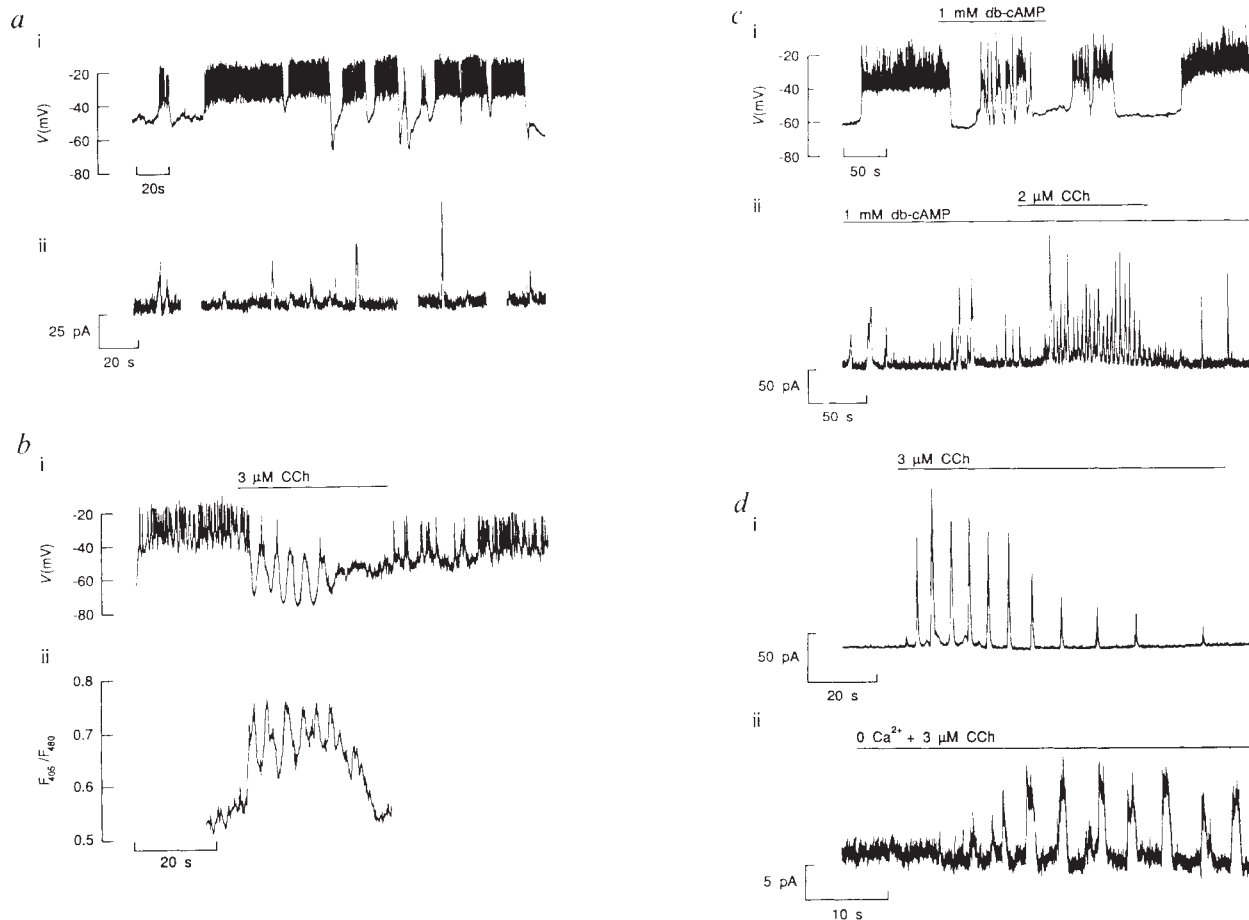


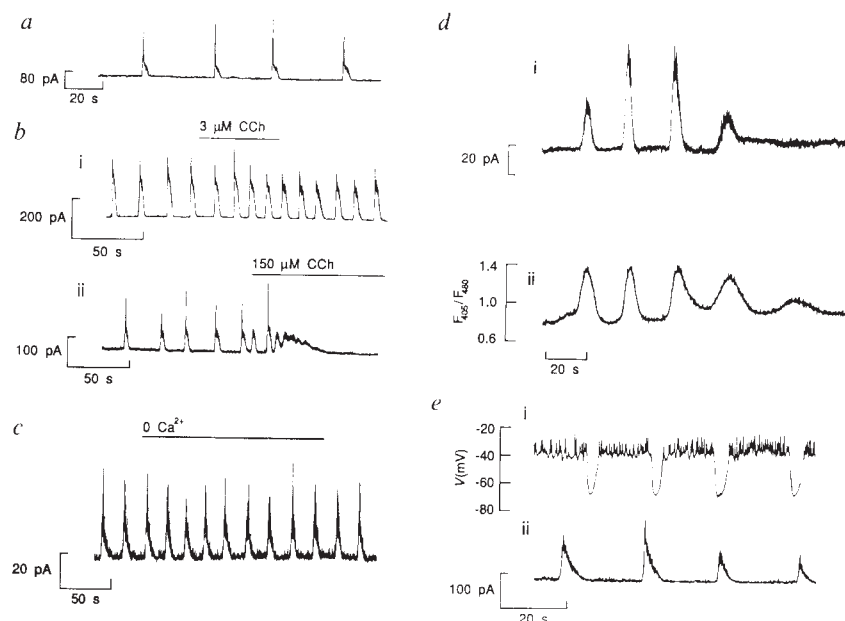
FIG. 1 Membrane current oscillations in intact β -cells. *a* (i), Oscillatory electrical activity recorded from a β -cell in a large cluster (estimated size 50–100 cells) stimulated with 10 mM glucose; (ii) spontaneous current fluctuations recorded from same β -cell under voltage-clamp conditions. During the periods indicated by the dotted horizontal lines, no oscillations were observed and the record was interrupted for 30–60 s to save space. *b* (i), Electrical activity recorded from β -cell in the presence of 20 mM glucose before and after addition of 3 μ M carbamylcholine; (ii) parallel recording of [Ca $^{2+}$] $_i$. The correlation between peak [Ca $^{2+}$] $_i$ and maximum repolarization is indicated by dotted vertical lines. *c* (i), Electrical activity recorded from β -cell in the presence of 10 mM glucose before, during and after exposure to 1 mM db-cAMP; (ii) membrane currents recorded from β -cell in the presence of 1 mM db-cAMP. Addition of 2 μ M carbamylcholine reversibly increases frequency of current oscillations. *d*, Membrane currents recorded in the presence of 10 mM glucose in response to stimulation with 3 μ M carbamylcholine in the presence (i) and absence (ii) of extracellular Ca $^{2+}$.

METHODS. Mouse pancreatic β -cells were isolated and maintained in tissue culture as previously described 29 . Whole-cell membrane currents were recorded using an EPC-7 patch-clamp amplifier (List Electronic). Unless otherwise indicated, the cells were held at 0 mV membrane potential, the

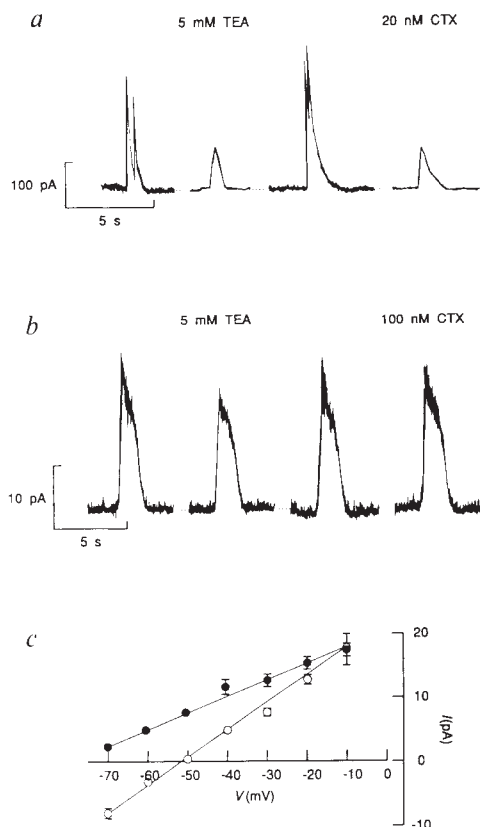
current signal filtered at 100 Hz (–3 dB) and cells bathed in an extracellular solution containing (in mM) 138 NaCl, 5.6 KCl, 1.2 MgCl $_2$, 2.6 CaCl $_2$ and 5 HEPES-NaOH (pH 7.4). In Figs 2 and 4, 0.5 mM tolbutamide was present to block ATP-regulated K $^+$ channels, which were not completely blocked by 3 mM intracellular ATP. Similar observations were made with tolbutamide-free media. Membrane potentials and currents were recorded using either the perforated patch whole-cell configuration as previously described 30 (intact cells; Fig. 1) or the conventional whole-cell technique (dialysis of cell interior). In *a* (ii), the currents were filtered at 20 Hz. In *d* (i), the extracellular solution was supplemented with 5 mM Co $^{2+}$. The Ca $^{2+}$ -free solution used in *d* (ii) consisted of (in mM) 138 NaCl, 5.6 KCl, 1.2 MgCl $_2$, 0.2 CoCl $_2$ and 5 HEPES-NaOH (pH 7.40). [Ca $^{2+}$] $_i$ was measured by dual emission spectrofluorimetry, using indo-1 (Molecular Probes) as the indicator and the hard- and software of Newcastle Photometric System adapted to a Zeiss Axiovert-10 microscope. [Ca $^{2+}$] $_i$ is quoted as the indo-1 fluorescence ratio F_{405}/F_{480} determined at 10 Hz. Excitation was effected at 360 nm. In *b* the β -cells were loaded with 1 μ M indo-1/AM (Molecular Probes) for 25 min before the experiments. During most of the experiments the β -cells were superfused (bath volume: 0.5 ml; rate: 4 ml min $^{-1}$). Experiments were performed at 27–30°C (Fig. 1) or at room temperature (20–22°C; Figs 2–4). Statistical significances were evaluated using Student's *t*-test of paired data.

FIG. 2 *a*, Spontaneous variations in whole-cell K^+ conductance in mouse β -cells evoked by inclusion of 50 μ M GTP γ S in the pipette (intracellular) solution. *b*, Effects of 3 μ M (i) and 150 μ M (ii) carbamylcholine on GTP γ S-induced membrane current oscillations. *c*, K^+ -current oscillations persist after removal of extracellular Ca^{2+} . *d*, Parallel recordings of membrane conductance (i) and $[Ca^{2+}]_i$ (ii) in the presence of 50 μ M GTP γ S. *e*, Membrane potential (i) and membrane currents (ii) recorded from β -cell dialysed with GTP γ S. (ii) is the direct continuation of (i) after switching the amplifier from the current-clamp to the voltage-clamp mode.

METHODS. In *a* and *b* the extracellular solution was supplemented with 5 mM Co^{2+} . Co^{2+} was omitted in *d* and *e* to avoid quenching of the indo-1 fluorescence or inhibition of spiking. In *c*, the medium was nominally Ca^{2+} free during the period indicated by the horizontal line. The intracellular solution used in Figs 2–4 consisted of (in mM) 140 KCl, 1 $MgCl_2$, 0.1 EGTA (or indo-1 (Molecular Probes) in panel *d*), 3 Mg -ATP and 5 HEPES-KOH (pH 7.15). The intracellular solution also contained 50 μ M GTP γ S (Boehringer). The holding potential in *e* (ii) was -20 mV.



an outward time-independent holding current. The latter observation suggest that a small fraction of the K_{ATP} channels remain active even in the presence of 3 mM ATP. The current (I)-voltage (V) properties of the TEA-resistant GTP γ S-induced current oscillations are summarized in Fig. 3*c*. With normal $[K^+]_o$, the TEA-resistant component displayed an extrapolated reversal potential of -77 ± 2 mV ($n=4$). After increasing $[K^+]_o$ to 20.6 mM, the zero current potential was -50 ± 3 mV ($n=6$). This shift of 27 mV is reasonably close to the 33 mV predicted by the Nernst equation for a 3.6-fold increase in $[K^+]_o$, indicating that the TEA-resistant current is selective for K^+ .



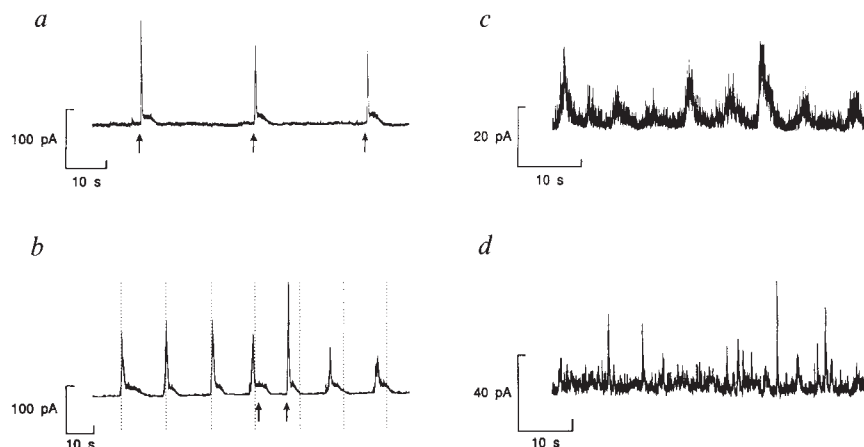
K^+ currents similar to those produced by infusion of GTP γ S could be induced by intracellular application of $InsP_3$. Figure 4*a* shows that applications of 15 μ M $InsP_3$, by release from a caged precursor¹⁹, evoked large K^+ currents even in the absence of exogenous GTP γ S. Photoreleased $InsP_3$ was unable to initiate the repetitive behaviour, although new K^+ -current oscillations could be elicited by subsequent $InsP_3$ additions. This argues that $InsP_3$ is rapidly degraded within the cell. The interaction of the effects of GTP γ S and $InsP_3$ were investigated. Whereas $InsP_3$ had only a marginal effect when applied at the end of a GTP γ S-induced oscillation (Fig. 4*b*, first arrow), a second application 10 s later (second arrow) evoked a large current. This suggests that GTP γ S evokes K^+ -current oscillations through an $InsP_3$ -dependent mechanism. As phospholipase C is known to be regulated by G proteins^{20,21}, it is conceivable that GTP γ S acts by stimulating $InsP_3$ formation. Infusion of β -cells with either 15 μ M of $InsP_3$ (Fig. 4*c*) or 10 μ M of the poorly metabolizable $InsP_3$ analogue²² $Ins(2,4,5)P_3$ (Fig. 4*d*) did not initiate the regular oscillatory pattern obtained with GTP γ S, but resulted in the appearance of continuous high frequency current fluctuations of variable amplitude. These findings indicate that simply elevating the cytoplasmic $InsP_3$ concentration is not sufficient to mimic the action of GTP γ S, raising the interesting possibility that variations of the $InsP_3$ concentration are critical²³.

We suggest that the oscillatory electrical activity recorded from β -cells in intact pancreatic islets at intermediate glucose concentrations results from a complex interplay between the β -cell and intraislet neurotransmitters and hormones (for example, acetylcholine, cholecystokinin and glucagon; see ref. 9 for review). It is implicit from this hypothesis that the failure to observe oscillatory electrical activity in most isolated tissue-cultured β -cells⁷ is attributable to lack of such interactions⁸.

FIG. 3 Pharmacological properties of GTP γ S-induced K^+ currents. *a*, Effects of extracellularly applied TEA (5 mM) and charybdotoxin (CTX, 20–100 nM; Alomone Labs) on K^+ conductance variations observed at 0 mV membrane potential. Effects of both compounds were reversible (not shown). TEA resulted in $85 \pm 5\%$ inhibition ($n=4$) and charybdotoxin produced a similar block ($60 \pm 8\%$; $n=3$). *b*, Effects of TEA and charybdotoxin at -40 mV. In the same cells, charybdotoxin and TEA produced $8 \pm 2\%$ and $23 \pm 2\%$ inhibition ($n=3$), respectively. *c*, I - V relationships recorded for the GTP γ S-induced current in the presence of 5 mM TEA at 5.6 mM $[K^+]_o$ (●) and after increasing $[K^+]_o$ to 20.6 mM (○). Data are quoted as mean values $\pm$ s.e.m. for seven (low $[K^+]_o$) and six (high $[K^+]_o$) experiments. The extracellular medium was supplemented with 5 mM Co^{2+} .

FIG. 4 InsP_3 -induced K^+ currents. *a*, InsP_3 ($15 \mu\text{M}$) was released three times from its caged precursor as indicated by the arrows in the absence of exogenous guanine nucleotides. *b*, GTP γS -induced ($50 \mu\text{M}$) membrane currents. $15 \mu\text{M}$ InsP_3 was released twice as indicated by arrows. The vertical dotted lines indicate the predicted positions of the GTP γS -induced peak currents using an interval of 17 s (mean of five oscillations: 17.4 ± 0.4 s). *c*, Currents observed during infusion of $15 \mu\text{M}$ InsP_3 through the low-resistance recording pipette. *d*, Currents observed in the presence of $10 \mu\text{M}$ $\text{Ins}(2,4,5)\text{P}_3$.

METHODS. The extracellular medium was supplemented with 5 mM Co^{2+} . Electrodes were filled with $20 \mu\text{M}$ caged InsP_3 (Calbiochem). Liberation was effected by 1 s of ultraviolet irradiation (300–450 nm). The efficiency of liberation was assumed to be the same as for caged ATP (70%), giving a concentration of InsP_3 of $\sim 15 \mu\text{M}$. The details of the photorelease apparatus are described elsewhere³¹. InsP_3 (K^+ -salt) was purchased from



Sigma. GTP γS was not added to the pipette solution used in panels *a*, *c* and *d*.

The InsP_3 -dependent Ca^{2+} -activated K^+ current we now describe constitutes a variable background current superimposed on the K^+ permeability due to K_{ATP} channel activity. At glucose concentrations above the threshold for insulin secretion, K_{ATP} -channel activity is substantially reduced, leading to membrane depolarization and generation of action potentials. When an extra oscillating K^+ current is added, however, the K^+ permeability will be sufficiently elevated to transiently repolarize the β -cell. At higher glucose concentrations, remaining K_{ATP} -channel activity is largely abolished²⁴. Under these conditions, the development of an extra K^+ current may be insufficient to repolarize the β -cell. Such a concept would also account for the observation that tolbutamide converts the oscillatory glucose-induced electrical activity observed at an intermediate glucose concentration into continuous spiking^{24,25}.

Many non-excitable cells exhibit InsP_3 -mediated periodic increases in $[\text{Ca}^{2+}]^{23}$ that are associated with changes in membrane conductance²⁶. We now provide evidence for such mechanisms being operational also in an excitable cell. It will be interesting to investigate whether processes similar to those present in the β -cell also contribute to oscillatory electrical activity in other cells (compare refs 27, 28). Indeed, this can easily be envisaged as both InsP_3 -sensitive Ca^{2+} stores and Ca^{2+} -activated K^+ channels are present in various excitable cells^{16,20,21}. □

27. Benham, C. D. & Bolton, T. B. *J. Physiol., Lond.* **381**, 385–406 (1986).
28. Nussinovitch, I. *J. Physiol., Lond.* **395**, 303–318 (1988).
29. Arkhammar, P., Nilsson, T., Rorsman, P. & Berggren, P.-O. *J. Biol. Chem.* **262**, 5448–5454 (1987).
30. Rorsman, P. et al. *Nature* **349**, 77–79 (1991).
31. Åmmälä, C., Bokvist, K., Galt, S. & Rorsman, P. *Biochim. biophys. Acta* **1092**, 347–349 (1991).

ACKNOWLEDGEMENTS. This work was supported by the Swedish Medical Research Council, the Bank of Sweden Tercentenary Foundation, the Swedish Diabetes Association, the Åke Wibergs Stiftelse, the Magn. Bergvalls Stiftelse, the Nordic Insulin foundation, the Swedish Hoechst, the Swedish Society for Medicine, the OE och Edla Johanssons Vetenskapliga Stiftelse and the Torsten och Ragnar Söderbergs Foundations. We thank R. F. Irvine (Cambridge, UK) for the gift of the stable $\text{Ins}(2,4,5)\text{P}_3$ analogue.

Precise prediction of a dominant class I MHC-restricted epitope of *Listeria monocytogenes*

Eric G. Pamer*†, John T. Harty* & Michael J. Bevan*

* Howard Hughes Medical Institute, Department of Immunology and
† Department of Medicine, University of Washington, Seattle,
Washington 98195, USA

Listeria monocytogenes is a Gram-positive bacterium which grows in the cytoplasm of eukaryotic cells and can cause severe disease in immunocompromised individuals^{1,2}. In murine systems CD8^+ T lymphocytes have been shown to be important effectors of acquired protective immunity against *L. monocytogenes*^{3–5}. Class I MHC-restricted CD8^+ cytotoxic T lymphocytes (CTL), which lyse J774 macrophage-like targets infected with *L. monocytogenes*, are induced following *in vivo* injection of live organisms. Natural peptide epitopes derived from *L. monocytogenes* can be acid-extracted from heavily infected BALB/c spleens and detected by CTL. A CTL clone, B9, derived from a (BALB/c $\times$ C57BL/6) F_1 ($\text{H}-2^{\text{dxb}}$) mouse, recognizes one of these natural epitopes in an $\text{H}-2\text{K}^{\text{d}}$ -restricted fashion. B9 also recognizes P815 ($\text{H}-2^{\text{d}}$) mastocytoma cells transfected with the listeriolysin gene. To identify the region of the listeriolysin recognized by CTL we used the $\text{H}-2\text{K}^{\text{d}}$ peptide-binding motif described by Rammensee and colleagues⁶ to synthesize 11 nonamer peptides. One of these peptides, listeriolysin 91–99, was recognized very efficiently by B9. This represents the first identified class I MHC-restricted epitope of bacteria and demonstrates the utility of the allele-specific motif for predicting CTL epitopes.

Strains of *L. monocytogenes* that express listeriolysin (LLO; relative molecular mass 59,000 (M_r 59K)) are virulent in mice and, when administered at a sublethal dose, induce protective immunity⁷. *L. monocytogenes* lacking LLO are avirulent and cannot induce immunity^{8–10}. (BALB/c $\times$ C57BL/6) F_1 mice were

Received 7 February; accepted 4 September 1991.

1. Henquin, J. C. & Meissner, H. P. *Experientia* **40**, 1043–1052 (1984).
2. Cook, D. L., Ikeuchi, M. & Fujimoto, W. Y. *Nature* **311**, 269–271 (1984).
3. Atwater, I., Rosario, L. M. & Rojas, E. *Cell Calcium* **4**, 451–461 (1983).
4. Kukuljan, M., Gonçalves, A. A. & Atwater, I. *J. membr. Biol.* **119**, 187–195 (1991).
5. Cook, D. L., Satin, L. S. & Hopkins, W. F. *Trends Neurosci.* **14**, 411–414 (1991).
6. Rosario, L. M., Atwater, I. & Scott, A. M. *Adv. exp. med. Biol.* **211**, 413–425 (1987).
7. Smith, P. A., Ashcroft, F. M. & Rorsman, P. *FEBS Lett.* **261**, 187–190 (1990).
8. Grapenjiesser, E., Gylfe, E. & Hellman, B. *J. Biol. Chem.* **266**, 12207–12210 (1991).
9. Rasmussen, H., Zawilich, K. C., Ganesan, S., Calle, R. & Zawilich, W. S. *Diabetes Care* **13**, 655–666 (1990).
10. Penner, E. & Neher, E. *J. exp. Biol.* **139**, 329–345 (1988).
11. Lund, P. E., Grapenjiesser, E., Gylfe, E. & Hellman, B. *Biochem. biophys. Res. Commun.* **177**, 777–783 (1991).
12. Capod, T., Noel, J., Combettes, L. & Claret, M. *Biochem. J.* **275**, 277–280 (1991).
13. Bokvist, K., Rorsman, P. & Smith, P. A. *J. Physiol., Lond.* **423**, 327–342 (1990).
14. Fotherazi, S. & Cook, D. L. *J. membr. Biol.* **120**, 105–114 (1991).
15. Miller, C., Moczydlowski, E., Latorre, R. & Phillips, M. *Nature* **313**, 316–318 (1985).
16. Blatz, A. L. & Magleby, K. L. *Nature* **323**, 718–720 (1986).
17. Lebrun, P., Atwater, I., Claret, M., Malaisse, W. J. & Herchuelz, A. *FEBS Lett.* **161**, 41–44 (1983).
18. Trube, G., Rorsman, P. & Ohno-Shosaku, T. *Pflügers Arch.* **407**, 493–499 (1986).
19. Walker, J. W., Somlyo, A. V., Goldman, Y. E., Somlyo, A. P. & Trentham, D. R. *Nature* **327**, 249–252 (1987).
20. Berridge, M. J. & Irvine, R. F. *Nature* **312**, 315–321 (1984).
21. Berridge, M. J. & Irvine, R. F. *Nature* **341**, 197–205 (1989).
22. Hill, T. D., Dean, N. M. & Boynton, A. L. *Science* **242**, 1176–1178 (1988).
23. Harootyanian, A. T., Kao, J. P. Y., Paranjape, S. & Tsien, R. Y. *Science* **251**, 75–78 (1991).
24. Henquin, J. C. *Biochem. biophys. Res. Commun.* **156**, 769–775 (1988).
25. Cook, D. L. & Ikeuchi, M. *Diabetes* **38**, 416–421 (1989).
26. Wakui, M., Potter, V. L. & Petersen, O. H. *Nature* **339**, 317–320 (1989).

infected with a sublethal dose of virulent *L. monocytogenes* (ATCC 43251) and eight days later splenocytes were taken and stimulated *in vitro* with J774 macrophage-like cells (H-2^d) infected with *L. monocytogenes*. After three cycles of stimulation the resulting T cells were >94% CD8⁺ with no detectable CD4⁺ cells (results not shown). One of these T cell lines, named LmT, lysed J774 cells only if they were infected with live, virulent *L. monocytogenes* (Fig. 1). J774 cells infected with avirulent *L. monocytogenes* that do not express LLO (ATCC 43250)¹¹ or treated with heat-killed virulent *L. monocytogenes* were not lysed significantly above background (Fig. 1). These findings agree with recent studies showing that the stimulation of γ -interferon production by *L. monocytogenes*-specific CD8⁺ T cells requires infection of presenting cells with live, virulent organisms¹².

To define the target antigen specificity of the LmT CTL line, spleens from BALB/c (H-2^d) mice infected with *L. monocytogenes* were extracted with 0.1% trifluoroacetic acid and material of less than M_r 5K was separated by reversed-phase high-pressure liquid chromatography (HPLC). Individual fractions were then tested for the presence of peptide antigens derived from *L. monocytogenes* by their ability to target CTL killing. LmT detected two main peaks (fractions 28 and 30) and one minor peak (fraction 35) of targeting activity on P815 (H-2^d) cells (Fig. 2a). This targeting activity is MHC-restricted as the

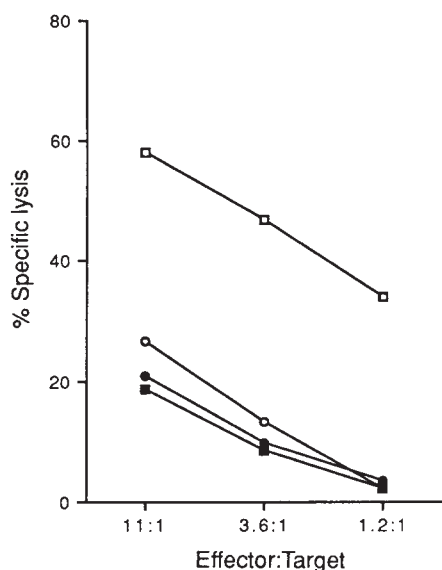


FIG. 1 CTL lysis of J774 cells infected with *L. monocytogenes*. J774 macrophages were treated with virulent *L. monocytogenes* (□), avirulent *L. monocytogenes* (○), heat-killed virulent *L. monocytogenes* (●) or left untreated (■). Effectors were added to ⁵¹Cr-labelled targets at the ratios indicated and percentage ⁵¹Cr release was determined as described²³.

METHODS. CTL line LmT was obtained from a (BALB/c × C57BL/6) F₁ mouse 8 days after infection with 5×10^3 virulent *L. monocytogenes* (ATCC 43251). Flasks for *in vitro* stimulation were prepared by adding 5×10^6 J774 cells in 10 ml RP10 lacking antibiotics²³. After 24 h the flasks were infected with 1 ml virulent *L. monocytogenes* in mid-log phase (A_{600} of 0.1) for 25 min, at which time the medium was replaced with RP10 containing $5 \mu\text{g ml}^{-1}$ gentamicin. Three hours later medium was replaced with RP10 containing 100 U ml^{-1} penicillin, $100 \mu\text{g ml}^{-1}$ streptomycin and $50 \mu\text{g ml}^{-1}$ gentamicin and incubated overnight. Immune splenocytes (3×10^7) were added to the flasks (in 20 ml volume) and incubated upright. CTL were restimulated every 7–12 days and after the first two stimulations the media was supplemented with interleukin-2. CTL were assayed by labelling 10^6 J774 cells with $100 \mu\text{Ci}$ ⁵¹Cr for 45 min in antibiotic-free media. Cells (10^4) were placed into wells and allowed to adhere for 30 min. Cells were then exposed to either 2×10^6 live or heat-killed (60 °C for 30 min) virulent or avirulent (ATCC 43250) *L. monocytogenes* for 25 min. Medium was replaced with RP10 containing $5 \mu\text{g ml}^{-1}$ gentamicin, CTL were added and lysis was allowed to proceed for 3 h. Spontaneous lysis of all target cells was <30%.

BALB/c trifluoroacetic acid extracts did not sensitize EL4 (H-2^b) cells and is *L. monocytogenes* dependent as uninfected BALB/c spleen extracts did not sensitize P815 targets. Additionally, trifluoroacetic acid extracts of infected C57BL/6 (H-2^b) spleens did not sensitize H-2^d targets. *L. monocytogenes*-specific CD8⁺ T lymphocytes that are not H-2 restricted have been described^{13,14}. Although CTL line LmT is H-2 restricted, we have generated other *L. monocytogenes*-specific CTL lines that recognize peptides extracted from infected allogeneic spleens in an H-2 unrestricted manner. The specificities of these CTL lines are being investigated.

CTL clone B9, which was derived from the LmT line by limiting dilution, is specific only for fraction 28 of the infected BALB/c spleen extract (Fig. 2b). H-2K^d-specific antibodies blocked lysis of P815 cells coated with fraction 28 whereas the H-2D^d-specific antibodies did not (Fig. 3a), indicating that the *L. monocytogenes*-derived peptide in fraction 28 is presented by H-2K^d. We were able to determine which bacterial protein

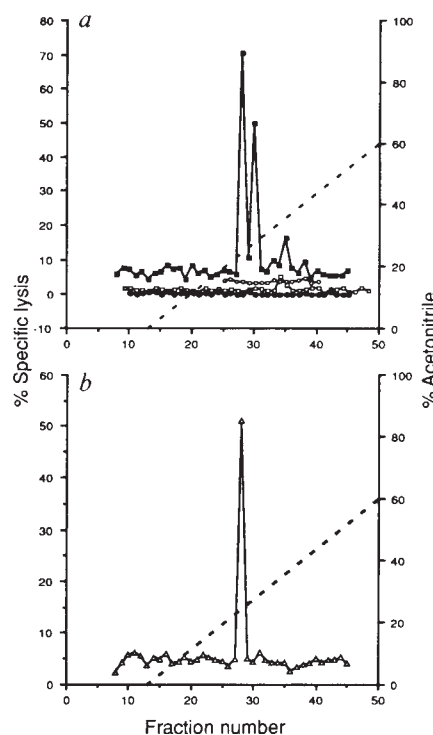
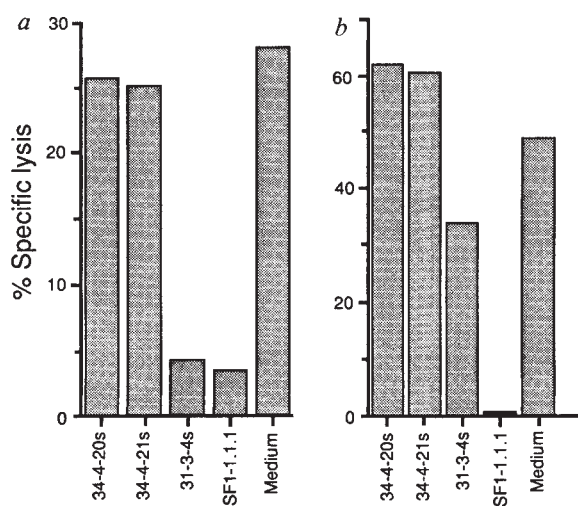


FIG. 2 Recognition by CTL line LmT and clone B9 of natural peptides derived from *L. monocytogenes* by acid extraction of infected spleens. a, CTL line LmT was used to lyse ⁵¹Cr-labelled P815 cells coated with reversed-phase HPLC fractions of acid extracts from infected BALB/c (■) and C57BL/6 (○) spleens and uninfected BALB/c (□) spleens. EL4 cells were also used with HPLC fractions from infected BALB/c spleens (●). Effector to target ratio was 20:1. b, CTL clone B9 was assayed on P815 cells coated with HPLC fractions of acid extracts from infected BALB/c spleens (▲). Effector to target ratio was 20:1.

METHODS. BALB/c mice were infected with 5×10^5 *L. monocytogenes* and C57BL/6 mice were infected with 10^6 *L. monocytogenes*. After 48 h the spleens were taken and homogenized sequentially with a tissue grinder and dounce homogenizer in 0.1% trifluoroacetic acid (TFA) and sonicated as described²⁴. This material was centrifuged at $100,000g$ for 30 min and the supernatant was passed over a Sephadex G-25 column. Material of less than 5K was collected and lyophilized. HPLC of the acid extract from one spleen in 0.1% TFA was performed with a C18 300A reversed-phase column using a 0–60% gradient of acetonitrile with 0.1% TFA at a rate of 1 ml min^{-1} and 1 ml fractions were collected. Fractions were lyophilized and suspended in $100 \mu\text{l}$ PBS. CTL assays were performed for 4 h with LmT and B9 effectors using either P815 or EL4 target cells in $200 \mu\text{l}$ wells, of which $50 \mu\text{l}$ represented the HPLC fraction.



provided this epitope as a P815 cell line transfected with the LLO gene (PHEM3) was lysed by LmT and the B9 CTL clone. Lysis of PHEM3 was also blocked by H-2K^d-specific antibodies (Fig. 3b).

The exact sequences of certain peptides that are derived from endogenously synthesized viral proteins and associate with class

FIG. 3 Presentation of HPLC Fraction 28 by H-2K^d and lysis of P815 cells transfected with LLO. **a**, P815 cells were coated with HPLC fraction 28 peptides and exposed to LmT CTL in the presence of monoclonal antibodies 34-4-20s and 34-4-21s, which are specific for H-2D^d, and 31-3-4s²⁵ and SF1-1.1.1, which are specific for H-2K^d. **b**, P815 cells transfected with LLO (PHEM3) were used as targets for LmT CTL in the presence and absence of anti-H-2^d antibodies as in **a**.

METHODS. P815 and PHEM3 cells were labelled with ⁵¹Cr and suspended in culture supernatants from hybridomas 34-4-20s, 34-4-21s, 31-3-4s, SF1-1.1.1 and control medium. Fraction 28 peptides were added to the P815 targets. LmT effectors were added to targets at a ratio of 20:1 and the specific lysis was determined after 3 h. PHEM3 was obtained by transfecting P815 cells with linearized pHAPr-1-neo²⁶ containing the entire LLO gene¹⁷. Bases 1-1,587 of the *hlyA* gene encoding LLO were cloned into the *Bam*HI site of the pHAPr-1-neo expression vector after polymerase chain reaction (PCR) amplification from *L. monocytogenes* DNA. The oligonucleotides used for priming were 5'-CCCGGATCCACCATGAAAAAATAATGCTAG-3' and 5'-GGATCCGGATCCTTATTCGATTGATTATC-3' which introduced flanking *Bam*HI sites. Electroporation of *Pvu*II linearized plasmid containing the insert into P815 cells and selection of G418 resistant clones was performed as described²³. P815 cells transfected with the vector without the insert were not lysed by the LmT CTL line (not shown).

I MHC molecules have recently been determined^{15,16}. Furthermore, sequencing the mixture of peptides associated with purified class I molecules has led to the prediction of MHC allele-specific peptide-binding motifs⁶. Thus, it has been proposed that H-2K^d-associated peptides are nine residues long and contain tyrosine at position 2 and either leucine or isoleucine

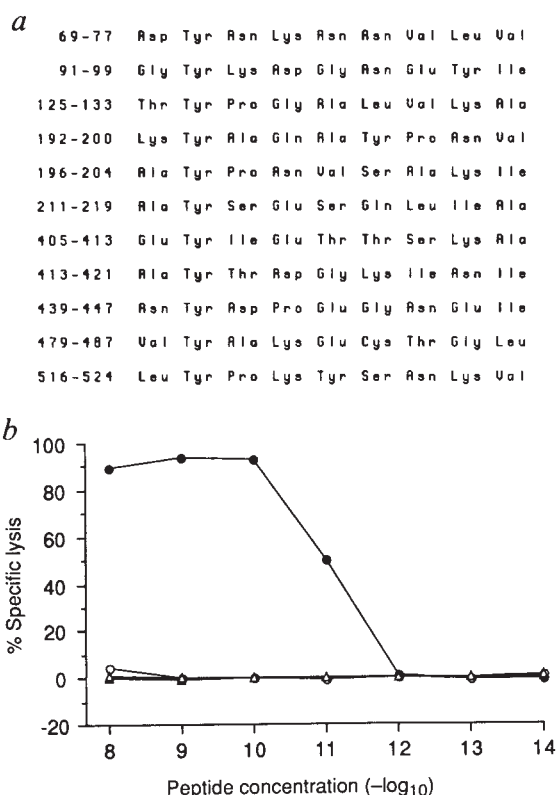
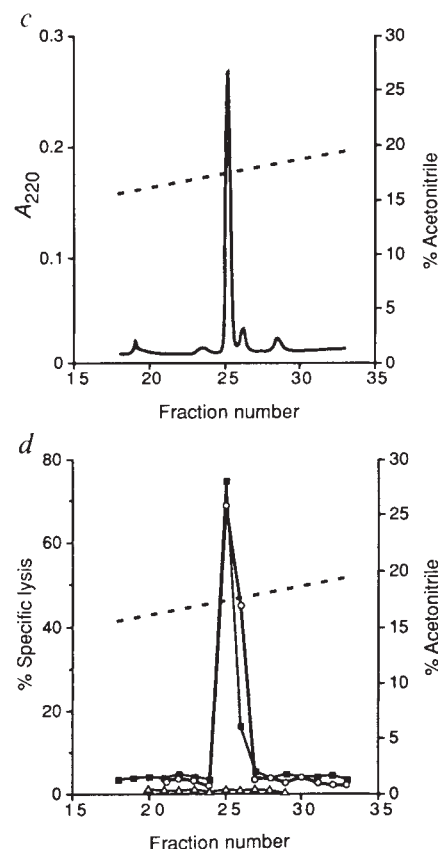


FIG. 4 Exact sequence determination of the H-2K^d-restricted LLO peptide. **a**, Eleven nonamer peptides that conform to the H-2K^d-binding motif were found in the LLO amino-acid sequence and were synthesized on resin with the F-moc multiple peptide synthesis method (Cambridge Research Biochemicals). The amino acids are numbered according to the sequence of LLO. **b**, All 11 peptides were used with P815 target cells and CTL assays were performed for 3 h with B9 CTL at an effector to target ratio of 20:1. LLO 91-99 sensitized targets (●) whereas LLO 196-204 (○), LLO 211-219



(△) and all other eight peptides (not shown) did not. **c**, HPLC A₂₂₀ profile of 2 µg synthetic LLO 91-99 (—) on a shallow acetonitrile gradient (broken line). **d**, CTL assays of shallow gradient HPLC fractions of 2 µg LLO 91-99 (each fraction diluted 10⁻⁶ for assay) (○), mock HPLC fractionation to assure that peptide contamination of the HPLC system had not occurred (△), and a TFA extract of a spleen from an *L. monocytogenes*-infected BALB/c mouse (■). B9 effectors were used to detect targeting activity on ⁵¹Cr-labelled P815 cells at a ratio of 10:1.

at the C terminus⁶. The predicted LLO sequence¹⁷ contains five such motifs (Fig. 4a). In addition, some identified foreign epitopes have valine, alanine or threonine at the C terminus. According to these criteria we synthesized 11 nonamers, representing only 18% of the LLO molecule, and tested them for their ability to sensitize P815 target cells to lysis by H-2K^d-restricted, LLO-specific CTL. Amino-acid residues 91-99 represent the H-2K^d LLO epitope recognized by B9 and this peptide can target cells at concentrations less than 10⁻¹¹ M (Fig. 4b). The precise molecular mass and partial amino-acid sequence of synthetic LLO 91-99 were confirmed by tandem ion spray mass spectroscopy. The other 10 peptides were not detected by LmT, suggesting that LLO contains only one H-2K^d-restricted peptide. The exact comigration of synthetic LLO 91-99 with the CTL targeting activity in trifluoroacetic acid extracts from infected BALB/c spleens on a very shallow acetonitrile gradient (Fig. 4c, d) indicates that we have identified the natural epitope. Other regions of the LLO molecule contain epitopes recognized by *L. monocytogenes*-specific CD4⁺ T lymphocytes¹⁸. LLO 91-99 was not, however, contained within any of the previously predicted T lymphocyte epitopes¹⁸.

Phagocytosed, virulent *L. monocytogenes* secrete LLO which disrupts the phagolysosomal membrane and allows the bacterium to enter the host cell cytoplasm¹. It has been proposed that only upon entry into the cytoplasm can *L. monocytogenes* antigens become processed and presented by class I MHC molecules¹². This would explain why *L. monocytogenes* strains or mutants lacking LLO do not induce protective immunity. Although our findings are consistent with this hypothesis, they could also be interpreted to suggest that the absence of LLO in avirulent *L. monocytogenes* is responsible for their lack of immunogenicity. Although the immunogenicity of LLO has been demonstrated^{18,19}, further studies will be required to settle this controversy.

Identification of the CTL epitopes of infectious agents is important for vaccine development²⁰⁻²². The diversity and number of proteins expressed by bacterial and protozoal pathogens has made the task of identifying CTL epitopes formidable. Determination of LLO 91-99 as a major H-2^d-restricted CTL epitope was possible because of the recent identification of the binding motif of H-2K^d for endogenous peptides⁶. Our findings demonstrate the predictive power of this motif for the determination of a pathogen-derived CTL epitope and argue the case for the systematic determination of binding motifs of other class I molecules. □

Received 24 July; accepted 10 September, 1991.

1. Tilney, L. G. & Portnoy, D. A. *J. Cell Biol.* **109**, 1597-1608 (1989).
2. Marget, W. & Seeliger, H. P. R. *Infection* **16**, 175-177 (1988).
3. Kaufmann, S. H. E., Hug, E. & Delibero, G. *J. exp. Med.* **164**, 363-368 (1986).
4. Bishop, D. K. & Hinrichs, D. J. *J. Immunol.* **139**, 2005-2009 (1987).
5. Mielke, M. E. A., Ehlers, S. & Hahn, H. *Infect. Immun.* **56**, 1920-1925 (1988).
6. Falk, K., Roetzschke, O., Stevanovic, S., Jung, G. & Rammensee, H.-G. *Nature* **351**, 290-296 (1991).
7. Mackaness, G. B. *J. exp. Med.* **116**, 381-406 (1962).
8. Gaillard, J. L., Berche, P. & Sansonetti, P. *Infect. Immun.* **52**, 50-55 (1986).
9. Kathariou, S., Metz, P., Hof, H. & Goebel, W. *J. Bact.* **169**, 1291-1297 (1987).
10. Portnoy, D. A., Jacks, P. S. & Hinrichs, D. J. *J. exp. Med.* **167**, 1459-1471 (1988).
11. Pine, L. *et al. J. clin. Microbiol.* **25**, 2247-2251 (1987).
12. Brunt, L. M., Portnoy, D. A. & Unanue, E. R. *J. Immunol.* **145**, 3540-3546 (1990).
13. Delibero, G. & Kaufmann, S. H. E. *J. Immunol.* **137**, 2688-2694 (1986).
14. Lukacs, K. & Kurlander, R. J. *J. Immunol.* **143**, 3731-3736 (1989).
15. Van Bleek, G. M. & Nathanson, S. G. *Nature* **348**, 213-216 (1990).
16. Roetzschke, O. *et al. Nature* **348**, 252-254 (1990).
17. Mengaud, J. *et al. Infect. Immun.* **56**, 766-772 (1988).
18. Safley, S. A., Cluff, C. W., Marshall, N. E. & Ziegler, H. K. *J. Immunol.* **146**, 3604-3616 (1991).
19. Berche, P., Gaillard, J., Geoffrey, C. & Alouf, J. E. *J. Immunol.* **139**, 3813-3821 (1987).
20. Deres, K., Schild, H., Wiesmueller, K., Jung, G. & Rammensee, H.-G. *Nature* **342**, 561-564 (1989).
21. Aichele, P., Hengartner, H., Zinkernagel, R. F. & Schulz, M. *J. exp. Med.* **171**, 1815-1820 (1990).
22. Kast, M. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 2283-2287 (1991).
23. Moore, M. W., Carbone, F. R. & Bevan, M. J. *Cell* **54**, 777-785 (1988).
24. Roetzschke, O., Falk, K., Wallny, H.-J., Faath, S. & Rammensee, H.-G. *Science* **249**, 283-287 (1990).
25. Ozato, K., Mayer, N. M. & Sachs, D. H. *Transplantation* **34**, 113-120 (1982).
26. Gunning, P., Leavitt, J., Muscat, G., Sun-Yu, N. & Keddes, L. *Proc. natn. Acad. Sci. U.S.A.* **84**, 4831-4835 (1987).

ACKNOWLEDGEMENTS. We thank S. Jameson for discussion. Supported by the National Institute of Allergy and Infectious Diseases of the NIH, and the Howard Hughes Medical Institute.

Phosphorylation and dephosphorylation of the high-affinity receptor for immunoglobulin E immediately after receptor engagement and disengagement

Rossella Paolini, Marie-Hélène Jouvin
& Jean-Pierre Kinet

Molecular Allergy and Immunology Section, National Institute of Allergy and Infectious Diseases, National Institutes of Health, Twinbrook II Bldg, 12441 Parklawn Drive, Rockville, Maryland 20852, USA

TRIGGERING of mast cells and basophils by immunoglobulin E (IgE) and antigen induces various biochemical signals, including tyrosine kinase activation^{1,2}, which lead to cell degranulation and the release of mediators of the allergic reaction. The high-affinity receptor for IgE (FcεRI) responsible for initiating these events is a complex structure composed of an IgE-binding α-chain, a β-chain and a homodimer of γ-chains³. It has been assumed that β and γ, which have extensive cytoplasmic domains, play an important but undefined role in coupling FcεRI to signal transduction mechanisms. Here we show that FcεRI engagement induces immediate *in vivo* phosphorylation on β (tyrosine and serine) and γ (tyrosine and threonine) by at least two different non-receptor kinases. We take advantage of unique features of this receptor system to demonstrate that the phosphorylation signal is restricted to activated receptors and is immediately reversible upon receptor

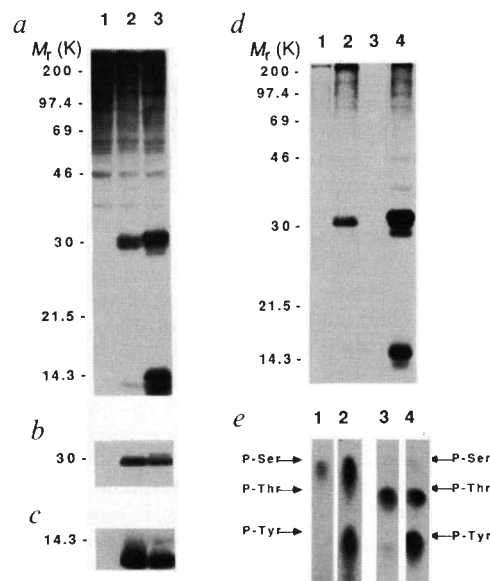
disengagement by undefined phosphatases. Rapid phosphorylation and dephosphorylation may be a general mechanism to couple and uncouple activated receptors to other effector molecules. This could be particularly relevant to other multimeric receptors containing FcεRI γ-chains or the related ζ and η chains such as the T-cell antigen receptor (TCR)⁴⁻⁶ and the low-affinity receptor for immunoglobulin G (FcγRIII, CD16) (refs 7-11).

Rat basophilic leukaemia cells (RBL-2H3) were incubated with anti-DNP IgE and labelled *in vivo* with [³²P]orthophosphate. The cells were then stimulated for one minute by a multivalent antigen dinitrophenyl[30]-human serum albumin (DNP[30]-HSA), solubilized and the IgE-bound receptors immunoprecipitated using a polyclonal anti-IgE antibody. Figure 1a shows the results of a reduced SDS-PAGE separation and autoradiography. In resting cells (lane 2), the β-chain appears as three species (relative molecular masses (*M_r*) around 30,000 (30K)) which are phosphorylated, although the fastest species requires a longer exposure to be clearly seen. After stimulation (lane 3), the overall level of phosphorylation of β increases substantially, particularly in the 33K species, although anti-β immunoblotting reveals equivalent amounts of β in immunoprecipitates from resting and stimulated cells (Fig. 1b). An increase in ³²P specific activity is also observed in the coprecipitating γ-chains, seen here in monomeric form of ~14K (Fig. 1a, c). Immunoprecipitations with monoclonal anti-β (ref. 12; Fig. 1d) and anti-γ-peptide antiserum⁶ (not shown) give virtually identical results. Thin-layer electrophoresis after gel extraction and hydrolysis of β- and γ-bands from similar experiments (Fig. 1e) confirms observations^{13,14} of serine phosphorylation of β and threonine phosphorylation of γ under resting conditions. But in contrast with previous studies^{13,14}, we find that receptor engagement induces tyrosine phosphorylation of both β and γ while increasing the serine phosphorylation of β and threonine phosphorylation of γ.

The direct relationship of receptor engagement and receptor phosphorylation was examined in antigen dose-response experiments (Fig. 2). The phosphorylation of β and γ increases with the dose of antigen and is maximal at 100 ng ml^{-1} of antigen. Receptor phosphorylation is not seen in ionomycin-stimulated cells (data not shown) and therefore is not simply secondary to nonspecific cell activation.

FIG. 1 Aggregation of Fc ϵ R1 in RBL-2H3 cells induces tyrosine and serine phosphorylation of β -chain as well as tyrosine and threonine phosphorylation of γ -chains. **a**, Immunoprecipitates with polyclonal anti-IgE antibody from ^{32}P -labelled cells were analysed by SDS-PAGE and autoradiography. M_r calibration is shown on the left. Cells were incubated with medium alone (lane 1) or with anti-DNP mouse IgE, then challenged for 1 min with medium alone (lane 2) or DNP[30]-HSA (lane 3). **b**, The same immunoprecipitates were analysed by western blotting with monoclonal anti- β antibody and **c**, with anti- γ -peptide antiserum. Immunoblots with irrelevant monoclonal antibody or preimmune serum did not show any reactivity (not shown). **d**, Immunoprecipitates with control IgG1 (lanes 1,3) or anti- β (lanes 2,4) antibody from labelled cells were analysed as in **a**. Cells were incubated with anti-DNP IgE and stimulated for 5 min with medium alone (lanes 1,2) or with DNP[30]-HSA (lanes 3,4). **e**, Thin-layer electrophoresis of phosphoaminoacids from non-activated receptor β -chains (lane 1) and γ -chains (lane 3) and from activated receptor β -chains (lane 2) and γ -chains (lane 4). **METHODS**. RBL-2H3 cells (1×10^7 cells per ml) were pre-incubated in phosphate-free medium for 4 h at 37°C before addition or not of $5 \mu\text{g ml}^{-1}$ of monomeric monoclonal anti-DNP mouse IgE (ref. 26) for 1 h at 37°C , then labelled at 3×10^7 cells per ml with $37 \text{ MBq per ml } [^{32}\text{P}]\text{-orthophosphate}$ (Amersham) for 2 h at 37°C . The cells (5×10^7 per sample) were incubated for various lengths of time with medium or with 100 ng ml^{-1} of DNP[30] conjugated to human serum albumin (Sigma). Cells were then immediately washed with 50 ml ice-cold PBS containing the following phosphatase inhibitors: 1 mM sodium vanadate, 5 mM sodium fluoride, 5 mM sodium pyrophosphate and 5 mM EDTA. Cells were lysed in a buffer (pH 8) containing 200 mM boric acid, 160 mM NaCl, 0.3% Triton X-100, 1% BSA, the same phosphatase inhibitors and the following protease inhibitors: $10 \mu\text{g ml}^{-1}$ aprotinin, $4 \mu\text{g ml}^{-1}$ leupeptin, $10 \mu\text{g ml}^{-1}$ pepstatin, 1 mM PMSF. The post-nuclear supernatants were precleared with protein A- or protein G-Sepharose beads (Pharmacia) and immunoprecipitated with rabbit polyclonal anti-IgE bound to protein A-Sepharose or with monoclonal anti- β (JRK) 12 or control IgG1 bound to protein G-Sepharose beads. Beads were washed with lysis buffer and the immunoprecipitates analysed on 12.5% SDS-PAGE under reducing conditions before autoradiography. For western blotting, the postnuclear supernatants from 1.5×10^7 cells were immunoprecipitated and analysed by SDS-PAGE as described and the proteins transferred electrophoretically to Immobilon-P membranes (Millipore) for 6 h at 50 V. Membranes were then incubated overnight in 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 5% BSA, before a two-hour incubation in the same buffer containing $15 \mu\text{g ml}^{-1}$ anti- β antibody or 1:200 dilution of anti- γ -peptide antiserum 6 or the same concentration of control antibody or pre-immune serum. After

We next investigated whether β - and γ -chain phosphorylations are early events in the signalling cascade, possibly preceding other early signals such as the hydrolysis of phosphoinositides and the increase in intracellular calcium, both of which are typically seen in RBL cells after 15 s of receptor engagement 15,16 . Figure 3a shows that, when the cells are saturated with IgE, phosphorylation of β and γ is almost maximal after



two washes in 20 mM Tris-HCl, pH 7.5, 500 mM NaCl, 0.1% Triton X-100, 2.5% BSA, membranes were incubated with 1:3,000 dilution of goat anti-mouse or goat anti-rabbit antibody conjugated to alkaline phosphatase (Jackson ImmunoResearch) and developed with the chemiluminescent substrate AMPPD $^{\text{TM}}$ (Tropix). Thin-layer electrophoresis of phosphoaminoacids was modified from ref. 27. Gel slices corresponding to β and γ were cut and the proteins digested with $50 \mu\text{g ml}^{-1}$ of trypsin in 50 mM NH_4HCO_3 to improve the yield of elution. The eluted peptides were hydrolysed for 90 min at 110°C in 0.2 ml 6N HCl. Samples were dissolved in 5 μl thin-layer buffer (10% glacial acetic acid, 1% pyridine in water, pH 3.5) containing 5 μg each of unlabelled phosphotyrosine, threonine, -serine (Sigma) as standards. Samples were applied to cellulose plates (100- μm DC-cellulose; Merck) and electrophoresed at 1,500 V for 45 min. Standards were developed by ninhydrin and ^{32}P -labelled amino acids identified by autoradiography. We have confirmed the results of this analysis by 2-dimensional thin-layer electrophoresis. Phosphoaminoacid analysis of β and γ purified from anti-IgE, anti- β or anti- γ immunoprecipitates gave virtually identical results.

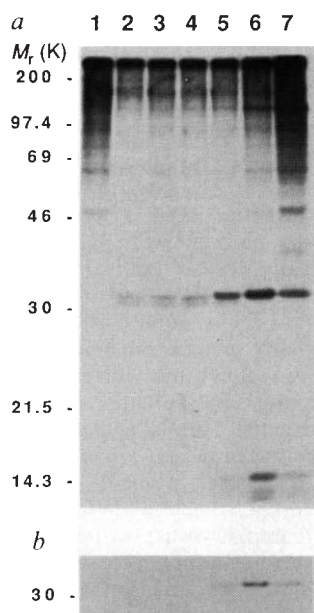


FIG. 2 Effect of increasing doses of antigen on β and γ phosphorylation. **a**, Immunoprecipitates with anti-IgE antibody from ^{32}P -labelled cells were analysed by SDS-PAGE and autoradiography as for Fig. 1. Cells were incubated with medium alone (lane 1) or with anti-DNP mouse IgE, then challenged for 5 min with medium alone (lane 2) or with 0.1 ng ml^{-1} (lane 3), 1 ng ml^{-1} (lane 4), 10 ng ml^{-1} (lane 5), 100 ng ml^{-1} (lane 6), $1 \mu\text{g ml}^{-1}$ (lane 7) of DNP-HSA. **b**, Tyrosine phosphorylation of β -chains from the same immunoprecipitates was analysed by immunoblotting with monoclonal anti-phosphotyrosine antibody 4G10 (UBI; ref. 28).

5 s stimulation with antigen (see legend). Indeed receptor phosphorylation is unaffected by the addition of EDTA at the time of stimulation, although [^3H]serotonin release is abolished under these conditions (not shown). In addition, efficient depletion of protein kinase C by prolonged incubation with phorbol esters did not modify receptor phosphorylation (not shown). Therefore it seems that the receptor-induced activation of tyrosine and serine/threonine kinases and the consequent receptor phosphorylation precede the other known biochemical signals.

One of the unique features of the IgE-Fc ϵ RI system is that it permits studies of rapid receptor disengagement. The binding of the monovalent ligand (IgE) is not sufficient for cell triggering. In fact, the β and γ from empty and liganded (but not cross-linked) receptors show similar states of phosphorylation when analysed by SDS-PAGE and subsequent thin-layer electrophoresis (not shown). To stimulate the cells, surface IgE-receptor complexes need to be engaged by a multivalent antigen. If a multihapten antigen is used as the triggering tool (here DNP[30]-HSA), then receptor disengagement can be studied by adding an excess of monovalent hapten (DNP-lysine). Figure 3a shows that within 5 s of monovalent hapten addition, β and γ are dephosphorylated, regardless of the length of the preceding triggering period (5, 15 or 60 s). The tyrosines, and to a lesser extent the serines, on β are dephosphorylated (Fig. 3b). Following disengagement by the addition of monovalent hapten, these receptors can then be rephosphorylated on β and γ (not shown)

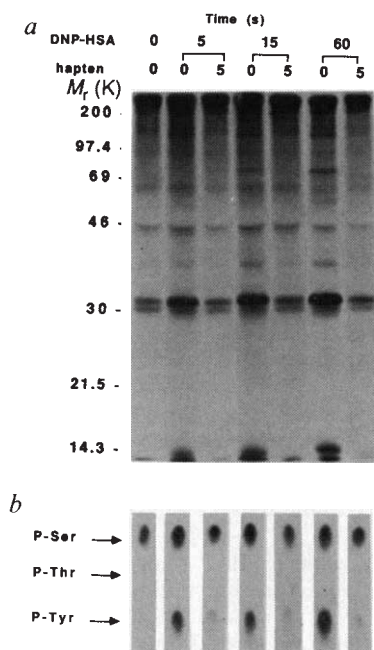


FIG. 3 Phosphorylation of β - and γ -chains is very rapid and reversible. *a*, Immunoprecipitates with monoclonal anti- β from ^{32}P -labelled cells were analysed by SDS-PAGE and autoradiography as for Fig. 1. Cells were incubated with IgE and stimulated with DNP-HSA for the time indicated. Cells were then chilled on ice or incubated for a further 5 s with 50 μM monovalent hapten (DNP-lysine). *b*, Thin-layer electrophoresis of phosphoaminoacids from receptor β chains purified from the corresponding lane of the gel shown above. Anti- β immunoblotting revealed an equivalent amount of β in the immunoprecipitates from this experiment (not shown). The levels of phosphorylation and dephosphorylation of β - and γ -chains were quantitated from similar experiments on a Betascope 603 blot analyser (Betagen). On average, the level of phosphorylation of β increases 2.5–5-fold after stimulation, whereas the level of phosphorylation of γ increases 100–400-fold. Both levels reach a maximum after 5–15 s. After disengagement with monovalent hapten, phosphorylation of β and γ decreases to resting levels. In control experiments (not shown), addition of monovalent hapten (DNP-lysine) had no effect on the phosphorylation of β and γ induced by monoclonal anti- α (BC4; ref. 17).

when re-engaged with anti-IgE. This rephosphorylation is qualitatively and quantitatively comparable to the phosphorylation of receptors freshly triggered with anti-IgE. Thus, phosphorylated receptors can be rapidly dephosphorylated by receptor disengagement and rephosphorylated following further stimulation.

Another advantage of this receptor system is the possibility to discriminate easily between activated and non-activated receptors in cells where only a fraction of receptors are bound to IgE and activated (Fig. 4). We analysed the fraction of receptors remaining empty (70%) by immunoprecipitation with monoclonal anti- α (BC4 recognizes only empty receptors)¹⁷. The level of phosphorylation in this fraction is the same whether the adjacent IgE-bound receptors (30%) are engaged (lane 7) or not (lane 6), even though an increase in phosphorylation is seen in the 30% fraction of activated receptors (lane 5). This experiment, in conjunction with anti- β (Fig. 4b) and anti- γ (Fig. 4c) immunoblots, demonstrates that the phosphorylation signal affects only β and γ associated with crosslinked receptors. This restricted phosphorylation may explain why phosphorylation is more intense when 100% (lane 3), as opposed to 30% (lane 5), of the receptors are activated.

Activation of tyrosine kinases is an early step in the signalling cascade of various receptors such as the TCR¹⁸, the B-cell antigen receptor¹⁹ and Fc γ RIII (CD16). As a consequence, receptor components such as the TCR- and CD16-associated ζ -chains^{20–22}, which are highly homologous with Fc ϵ RI

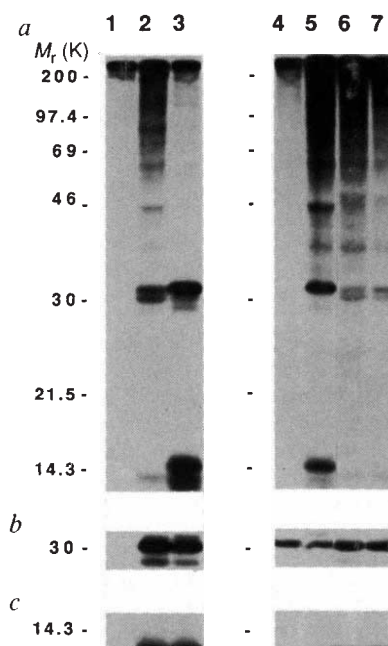


FIG. 4 Phosphorylation of β - and γ -chains is restricted to activated receptors. *a*, Immunoprecipitates from ^{32}P -labelled cells were analysed by SDS-PAGE and autoradiography as in Fig. 1. Lanes 1–5, anti-IgE immunoprecipitates: lane 1, cells unoccupied with IgE; lane 2, cells saturated with IgE and non-triggered; lane 3, cells saturated with IgE and triggered with antigen; lane 4, cells partially saturated and non-triggered; lane 5, cells partially saturated with IgE and triggered with antigen; lanes 6 and 7, immunoprecipitation with monoclonal anti- α BC4 (ref. 17) following a first immunoprecipitation with anti-IgE: lane 6, cells partially saturated and non-triggered; lane 7, cells partially saturated with IgE and triggered with antigen. *b*, The same immunoprecipitates analysed by western blotting with monoclonal anti- β antibody, and *c*, with anti- γ -peptide antiserum.

METHODS. RBL cells were incubated with medium alone (lane 1) or with saturating amount of IgE (10 $\mu\text{g ml}^{-1}/1 \times 10^7$ cells) (lanes 2, 3) or with a concentration of IgE (0.2 $\mu\text{g ml}^{-1}/1 \times 10^7$ cells) calculated to occupy about 20–30% receptors (lanes 4–7). Cells were challenged for 1 min with medium alone (lanes 1, 2, 4, 6) or with 100 ng ml^{-1} DNP-HSA (lanes 3, 5, 7).

γ -chains and clearly essential for cell activation²³, are phosphorylated on tyrosine. Indeed the genistein-induced inhibition of histamine release by RBL cells suggests a physiologically relevant role for tyrosine kinases in the Fc ϵ RI receptor activation pathway²⁴. Our results also support an important role for tyrosine kinase activity in receptor signalling. In addition, we have demonstrated that receptor phosphorylation, at least of Fc ϵ RI, is restricted to activated receptors, requires continuous receptor engagement and is immediately reversible upon receptor disengagement, presumably through the action of undefined phosphatases. We also show that Fc ϵ RI-induced activation of serine/threonine kinase(s) seems to occur as rapidly as tyrosine kinase activation. We propose that cycles of kinase and phosphatase activation are initiated and maintained by receptor engagement. The consequent phosphorylation/dephosphorylation of receptor and other substrates may promote coupling/uncoupling to other components of the signalling machinery. This model would explain the absence of an increase in phosphorylation over time, as well as the immediate receptor dephosphorylation that occurs when receptor triggering is halted.

The identification of the tyrosine and serine/threonine kinases involved in Fc ϵ RI activation will further our understanding of the system. Preliminary studies²⁵ suggest that the *lyn* and *yes* proteins of the *src* family of tyrosine kinases may be activated by Fc ϵ RI. Mutation analysis of specific residues on β and γ should enable us to assess the relative importance of tyrosine and serine/threonine receptor phosphorylations and to analyse their potentially discreet roles in coupling Fc ϵ RI to the signalling machinery in mast cells. □

Received 28 June; accepted 3 September 1991.

1. Benhamou, M., Gutkind, J. S., Robbins, K. C. & Siraganian, R. P. *Proc. natn. Acad. Sci. U.S.A.* **87**, 5327-5330 (1990).
2. Connelly, P. A., Farrell, C. A., Merenda, J. M., Conklyn, M. J. & Showell, H. J. *Biochem. biophys. Res. Commun.* **177**, 192-201 (1991).
3. Blank, U. *et al. Nature* **337**, 187-189 (1989).
4. Weissman, A. M. *et al. Science* **239**, 1018-1021 (1988).
5. Jin, Y. J. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 3319-3323 (1990).
6. Orloff, D. G., Ra, C., Frank, S. J., Klausner, R. D. & Kinet, J.-P. *Nature* **347**, 189-191 (1990).
7. Ra, C., Jouvin, M.-H. E., Blank, U. & Kinet, J.-P. *Nature* **341**, 752-754 (1989).
8. Hibbs, M. L. *et al. Science* **246**, 1608-1611 (1989).
9. Kurosaki, T. & Ravetch, J. V. *Nature* **342**, 805-807 (1989).
10. Anderson, P. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 2274-2278 (1990).
11. Lanier, L. L., Yu, G. & Phillips, J. H. *Nature* **342**, 803-804 (1989).
12. Rivera, J., Kinet, J.-P., Kim, J., Pucillo, C. & Metzger, H. *Molec. Immun.* **25**, 647-661 (1988).
13. Perez-Montfort, R., Fewtrell, C. & Metzger, H. *Biochemistry* **22**, 5733-5737 (1983).
14. Quarto, R. & Metzger, H. *Molec. Immun.* **23**, 1215-1223 (1986).
15. Millard, P. J., Gross, D., Webb, W. W. & Fewtrell, C. *Proc. natn. Acad. Sci. U.S.A.* **85**, 1854-1858 (1988).
16. Cunha-Melo, J. R., Dean, N. M., Moyer, J. D., Maeyama, K. & Beaven, M. A. *J. biol. Chem.* **262**, 11455-11463 (1987).
17. Basciano, L. K., Berenstein, E. H., Kmak, L. & Siraganian, R. P. *J. biol. Chem.* **261**, 11823-11831 (1986).
18. Klausner, R. D. & Samelson, L. E. *Cell* **64**, 875-878 (1991).
19. Lane, P. J. L. *et al. J. Immunol.* **146**, 715-722 (1991).
20. Baniyash, M., Garcia-Morales, P., Luong, E., Samelson, L. E. & Klausner, R. D. *J. biol. Chem.* **263**, 18225-18230 (1988).
21. O'Shea, J. J., Weissman, A. M., Kennedy, I. C. S. & Ortaldo, J. R. *Proc. natn. Acad. Sci. U.S.A.* **88**, 350-354 (1991).
22. Vivier, E. *et al. J. Immun.* **146**, 206-210 (1991).
23. Irving, B. A. & Weiss, A. *Cell* **64**, 891-901 (1991).
24. Stephan, V. *et al. FASEB J.* **5**, abst. 7578 (1991).
25. Eiseman, E. & Bolen, J. *Cancer Cells* **2**, 303-310 (1990).
26. Liu, F. T. *et al. J. Immun.* **124**, 2728-2737 (1980).
27. Cooper, J. A., Sefton, B. M. & Hunter, T. *Meth. Enzym.* **99**, 387-402 (1983).
28. Druker, B. J., Manon, H. J. & Roberts, T. M. *New Engl. J. Med.* **321**, 1383-1391 (1989).

ACKNOWLEDGEMENTS. We thank D. Ferris for advice in setting up *in vivo* phosphorylation experiments, R. Numerof for reviewing the manuscript, R. Siraganian for BC4 antibody and B. Druker for 4G10 antibody.

Interleukin-2 programs mouse $\alpha\beta$ T lymphocytes for apoptosis

Michael J. Lenardo

Laboratory of Immunology, Bldg 10, Rm 11N311, National Institute for Allergy and Infectious Diseases, National Institutes of Health, Bethesda, Maryland 20892, USA

ANTIGEN receptor stimulation of mature $\alpha\beta$ T lymphocytes can lead either to proliferation or death¹⁻⁴. Programmed cell death, termed apoptosis, leads to the clonal deletion of both thymocytes and mature T cells that establishes tolerance⁵⁻⁹. How a mature T cell selects between proliferation and death is not understood. Here I show that interleukin-2 (IL-2) is a critical determinant of the choice between these two fates. Both CD4⁺ and CD8⁺ T cells previously exposed to IL-2 undergo apoptosis after antigen-receptor stimulation. Antibody blockade of IL-2 but not IL-4 reverses the marked reduction of lymph node V β 8⁺ T cells caused in mice by the bacterial superantigen *Staphylococcus aureus* enterotoxin B. IL-2 may thus participate in a feedback regulatory mechanism by predisposing mature T lymphocytes to apoptosis.

A.E7 is a CD4⁺, T_H1 T lymphocyte clone that constitutively expresses high-affinity IL-2 receptors and shows increased thymidine incorporation after application of 100 units per ml of exogenous IL-2 for two days^{10,11}. Antigen stimulation at 1 μ M following IL-2 treatment diminished thymidine incorporation, possibly because of an antigen-dependent block in proliferation^{12,13}. Microscopic examination unexpectedly revealed extensive death of the T cells (Fig. 1a, top panels). There were 82% fewer IL-2 pre-treated T cells following 1 μ M antigen stimulation compared with control (Table 1). Cell death was less at lower doses of IL-2 and just evident without IL-2 pre-treatment, which could be attributed to IL-2 produced by antigen stimulation (Table 1). These results suggested the possibility that IL-2 following antigen stimulation leads to proliferation, whereas IL-2 before antigen exposure causes cell death.

I therefore evaluated IL-2 and antigen stimulation in an experiment in which endogenous IL-2 was not produced. A.E7 cells require co-stimulatory signals from antigen-presenting cells in addition to T-cell receptor stimulation to produce IL-2 (refs 14, 15). Consequently, A.E7 cells were pre-treated with IL-2 and stimulated without antigen-presenting cells on dishes coated with a monoclonal antibody against CD3 ϵ (ref. 16). Anti-CD3 ϵ stimulation following IL-2 again led to T-cell death (Fig. 1a, lower panels), with 74% of the untreated cells, but only 14% of the T cells pre-treated with 10 units per ml of IL-2, surviving (Table 1). Killing caused by anti-CD3 ϵ was also dependent on the IL-2 dose, with 49% cell loss at 2 units per ml. DNA fragments of 200-base-pair multiples were evident in dying A.E7 cells after IL-2 and anti-CD3 ϵ stimulation (Fig. 1b, 2C11) and death was avoided by including cyclosporin A as described in Table 1 legend. These data strongly suggested that apoptosis was occurring^{1,2,5,7}.

I then tested whether IL-2 could predispose cells bearing particular T-cell receptors (and not bystander cells) to apoptosis in a heterogeneous lymph node T-cell (LNT) population. Because LNT cells do not constitutively express IL-2 receptors, they were first stimulated to express high-affinity IL-2 receptors with the lectin, concanavalin A. The concanavalin A was then removed and the cells exposed to IL-2. As LNT cells did not survive without any IL-2, low doses (3 units per ml) and high doses (100 units per ml) of IL-2 were compared. After two days exposure to IL-2, the LNT cells (>97% $\alpha\beta$ T cells; data not shown) were plated on dishes coated with either no antibody, the F23.1 monoclonal antibody (anti-V β 8, specific for the V β 8.1, 2, 3 receptor chains)¹⁷ or the RR4-7 monoclonal antibody (anti-V β 6)¹⁸. In low IL-2, both anti-V β 8 and anti-V β 6 antibodies caused cell number to increase, but high IL-2 caused loss of cells (Table 1). Flow cytometry revealed that anti-V β 8 antibody markedly deleted cells with cognate V β 8 receptors in LNT cells given high IL-2 but not those given low IL-2 (data not shown). The high-dose IL-2 populations were gated separately into CD4⁺ cells and CD4⁻ cells (virtually all CD8⁺ cells), which revealed that anti-V β 8 decreased V β 8⁺ cells from 38.4% to 14.1% for

TABLE 1 Effect of IL-2 and antigen-receptor stimulation on T-cell viability

Cells	Experi- ment	Pretreatment	Cell number per well ($\times 10^{-4}$)		Stimulated	
			Control	1 μ M antigen	Control (%)	
A.E7	1	No IL-2	3.8 $\pm$ 0.3	2.3 $\pm$ 1.8	60	
		100 units IL-2	4.9 $\pm$ 2.5	0.9 $\pm$ 0.6	18	
	2	No IL-2	3.6 $\pm$ 0.4	2.8 $\pm$ 0.6	78	
		2 units IL-2	5.0 $\pm$ 0.4	2.4 $\pm$ 0.1	48	
		5 units IL-2	6.2 $\pm$ 0.5	2.4 $\pm$ 0.4	39	
		10 units IL-2	6.8 $\pm$ 0.6	2.0 $\pm$ 0.2	29	
		50 units IL-2	6.8 $\pm$ 1.0	2.0 $\pm$ 0.7	29	
	3	No IL-2	Control	Anti-CD3 ϵ 10 μ g ml $^{-1}$		
			4.3 $\pm$ 0.5	3.2 $\pm$ 0.4	74	
			2 units IL-2	7.1 $\pm$ 1.5	3.6 $\pm$ 0.1	51
			5 units IL-2	6.4 $\pm$ 0.8	2.1 $\pm$ 0.6	33
			10 units IL-2	7.4 $\pm$ 1.5	1.0 $\pm$ 0.2	14
	4	No IL-2	Anti-CD3 ϵ +CsA		—	
			3.9 $\pm$ 0.5	4.4 $\pm$ 0.5		
			25 units IL-2	2.2 $\pm$ 0.5	4.2 $\pm$ 0.9	—
LNT			Control	Anti-V β 8 20 μ g ml $^{-1}$		
			3 units IL-2	5.7 $\pm$ 1.1	9.1 $\pm$ 1.2	160
		100 units IL-2	Control	Anti-V β 6 33 μ g ml $^{-1}$		
			3 units IL-2	9.4 $\pm$ 1.0	165	
		100 units IL-2	Control	26.8 $\pm$ 4.7	67	

Cell counts ($\times 10^{-4}$) are averages of 4–6 independent haemocytometer counts of three wells determining only trypan-blue-excluding cells. Antigen was the amino acids 81–104 peptide from pigeon cytochrome *c* (gift from B. Beverly). The anti-CD3 ϵ antibody 145-2C11 (ref. 16) was used at the concentrations indicated, except in experiment 4, where 2.5 μ g ml $^{-1}$ was used. Plates were coated with 20 μ g ml $^{-1}$ anti-V β 8 (F23.1 antibody) and 33 μ g ml $^{-1}$ anti-V β 6 antibody (RR4-7 antibody) as described in the legend to Fig. 1. Control experiments in which equivalent amounts of antibody recognizing CD4, MHC class I, or CD45 were coated on plates had no effect on cell viability (data not shown). Cells were incubated in dishes for 48 h. Trypan blue-stained cells made up 30–70% of the differences between stimulated and controls. Cyclosporin A (gift from Sandoz) was included at 100 ng ml $^{-1}$ only during the stimulation by 145-2C11 antibody. IL-2 was human recombinant IL-2 (from C. Reynolds) or supernatant from MLA-144 cells (from the Fermentation Laboratory, FCRF, NCI), both of which gave essentially identical results. Data represent 8 experiments.

CD4 $^{+}$ cells, and from 38.0% to 19.4% for CD4 $^{-}$ (CD8) cells (Table 2). V β 6 $^{+}$ cells seem to be relatively increased to compensate for the loss of V β 8 $^{+}$ cells (Table 2). Similarly, anti-V β 6 deleted V β 6 $^{+}$ cells (from 12.3% to 1% for CD4 $^{+}$ cells, and from 10.0% to 2.2% for CD4 $^{-}$ (CD8) cells) but not V β 8 $^{+}$ cells (Table 2). These findings were not due to T-cell receptor downmodulation because: (1) apoptosis and decreased cell numbers were observed; (2) cells bearing heterologous receptors were relatively increased; and (3) no T-cell-receptor-negative cells were detected (data not shown). Thus, IL-2 predisposes to an endogenous death pathway in both CD4 $^{+}$ and CD8 $^{+}$ T cells. Bystander cells, although competent to undergo apoptosis, are unaffected by antigen-receptor-mediated killing of a LNT subpopulation.

If IL-2 preceding antigen receptor occupancy leads to apoptosis, then repetitive immunization could lead to antigen restimulation of cells exposed to IL-2 and cause IL-2-dependent deletion of T cell clones *in vivo*. To test this, BALB/c mice were given *Staphylococcus aureus* enterotoxin B (SEB) using an intravenous loading dose of 500 μ g followed by two injections of 125 μ g at two-day intervals. After eight days, peripheral lymph node T cells were analysed for V β 8 $^{+}$ cells (specifically activated by SEB) and V β 6 $^{+}$ cells (not SEB-reactive) (Fig. 2)¹⁹. In an

TABLE 2 Flow cytometric quantitation of *in vitro* deletion of V β 8- and V β 6-bearing LNT cells using anti-receptor antibodies

Gating	Fraction of total gated cells positive for:	Stimulation after IL-2 pre-treatment		
		None	Anti-V β 8	Anti-V β 6
CD4 $^{+}$ cells	V β 8	38.4%	14.1%	40.3%
	V β 6	12.3%	15.9%	1.0%
CD4 $^{-}$ (CD8) cells	V β 8	38.0%	19.4%	40.6%
	V β 6	10.0%	16.4%	2.2%

Lymph node T cells pretreated with 100 units per ml IL-2 were prepared as described in the legend to Fig. 2 and stimulated on culture dishes coated with either no antibody or monoclonal antibody raised against either V β 8 (F23.1) or V β 6 (RR4-7). Cells recovered from the plates were stained for two-colour cytometry with both anti-CD4 (Becton Dickinson) and either anti-V β 8 (F23.1) or anti-V β 6 (RR4-7). The CD4 staining was used to gate the cells into CD4 $^{+}$ and CD4 $^{-}$ pools; control staining showed that virtually all CD4 $^{-}$ cells were CD8 $^{+}$ (using anti-Lyt2) and cells were >97% $\alpha\beta$ T cells (using H57-597) in these preparations. The gated pools were then quantitated by flow cytometry using a Becton Dickinson FACSCAN for the fraction of either V β 8 $^{+}$ or V β 6 $^{+}$ cells. Independent gating was necessary for accurate quantitation because of the overgrowth of CD8 cells in antibody-stimulated samples²⁹. The fraction of the gated pool that was positive for either V β 8 or V β 6 is given as per cent; values in bold typeface are for conditions in which deletion was observed. Data represent 5 experiments.

uninjected animal, 22.3% of the T cells were detected by the antibody KJ16-33 (ref. 20), which recognizes V β 8.1, 8.2 receptors. As predicted, repetitive SEB immunization reduced the V β 8.1, 8.2 $^{+}$ cells to 7.5%, which is a decrease of >60%. Injection of SEB together with 800 μ g 3C7, an IL-2-receptor α -chain blocking antibody^{21,22} (anti-IL-2R; a gift from A. Kruisbeek), every 12 h intraperitoneally strikingly reversed the loss of V β 8.1, 8.2 $^{+}$ cells to 18.1%. Co-injection of antibody 11B11, which blocks IL-4 responses *in vivo*^{23,24} (anti-IL-4; a gift from W. Paul), did not reverse the loss of V β 8.1, 8.2 $^{+}$ cells caused by SEB. Injection of 3C7 or 11B11 antibody alone had no effect (data not shown). The fractions (mean $\pm$ s.d.) of V β 8.1, 8.2 $^{+}$ cells for several mice were: normal, 23.3 $\pm$ 0.6% (n = 4); SEB only, 9.7 $\pm$ 2.8% (n = 4); SEB + anti-IL-2R, 19.5 $\pm$ 3.8% (n = 3); and SEB + anti-IL-4, 9.4 $\pm$ 1.8% (n = 3). On day 2 after immunization, the fraction of V β 8 $^{+}$ cells was expanded in both SEB-only and the (SEB + IL-2R) mice, indicating that V β 8 $^{+}$ cells are not completely rescued by anti-IL-2R, possibly because of incomplete receptor blockade (data not shown). V β 6 $^{+}$ T cells were not deleted in these mice. Thus, clonal elimination under these conditions depends on IL-2 but not IL-4.

T lymphocyte apoptosis induced by IL-2 and antigen receptor stimulation has features of feedback inhibition²⁵: (1) it is caused by an 'end-product', for example IL-2, of the initial antigen stimulation; (2) apoptosis is greater with increasing doses of IL-2; and (3) it reverses the increase in the numbers of T cells initially caused by antigen^{1,4,8,9}. T-cell clonal specificity is maintained by the requirement for antigen stimulation, but antigen receptor occupancy alone is insufficient for apoptosis. This feedback pathway could be called *propiocidal* regulation (from Latin: *proprius*, "one's own") to indicate selective killing of the stimulated T cells, their progeny, and clones of related specificity. The lethality of *Staphylococcus* enterotoxins, in particular, seems to be due to substances produced by activated T cells^{19,26}. IL-2-mediated apoptosis could eliminate the affected T cells and decrease the harmful effects of chronic exposure to these toxins.

My findings implicate IL-2 in the clonal elimination of mature T cells, a mechanism postulated for extra-thymic tolerance^{1,4,8,9}. It has been shown that deletion of V β 8 $^{+}$ T cells in lymph node and spleen associated with apoptosis *in situ* following SEB injection¹. I find faster and greater loss of V β 8 $^{+}$ T cells when

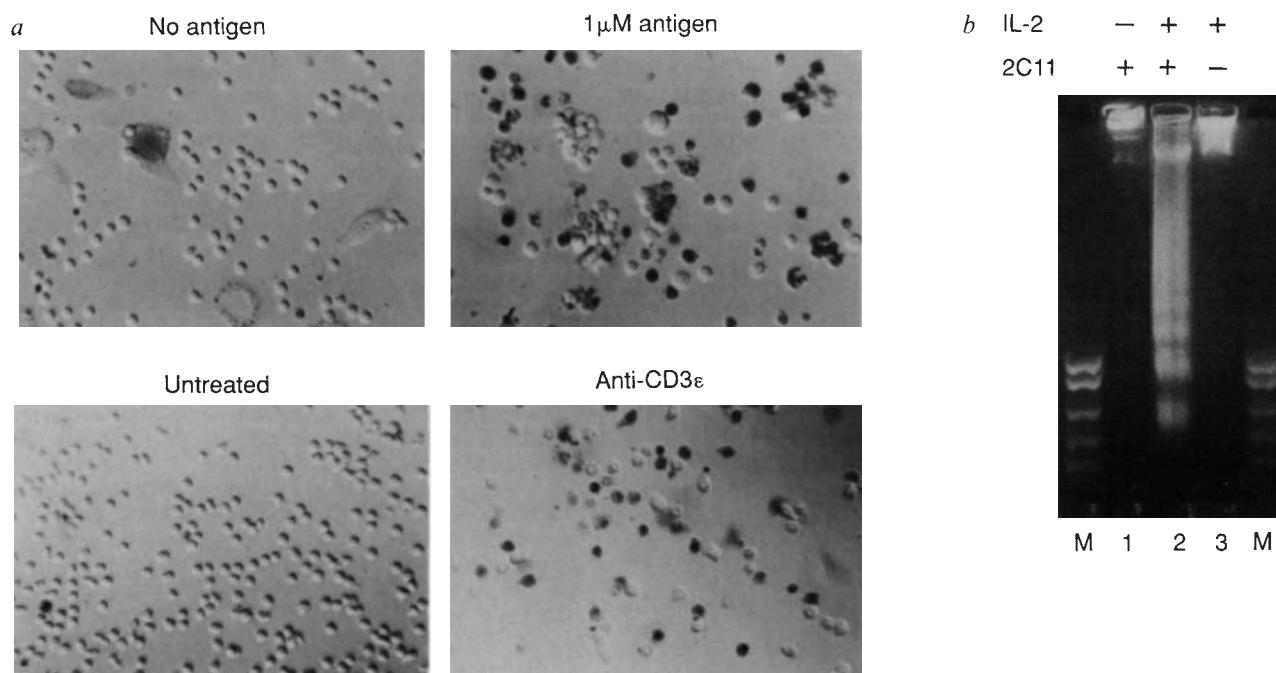


FIG. 1 Apoptosis resulting from antigen or anti-CD ϵ stimulation of A.E7 T cells after IL-2 pretreatment. *a*, Photomicrographs of A.E7 cells pretreated with 100 units of IL-2, stimulated and stained with trypan blue. Top panels, A.E7 cells (small round cells) and DCEK antigen-presenting cells (large fibroblastic cells) with no antigen (left) or 1 μ M antigen (right). Note that antigen-activated A.E7 cells also lyse DCEK antigen-presenting cells. Bottom panels, A.E7 cells in untreated wells (left) or wells precoated with 10 μ g ml $^{-1}$ anti-CD3 ϵ (right). Some nonadherent dead T cells and cell fragments are washed away during the staining, otherwise dead cells stain darkly. *b*, Agarose gel electrophoresis of DNA from A.E7 T cells treated with IL-2 and/or anti-CD ϵ (145-2C11). Lane M shows DNA size markers.

METHODS. A.E7 is a nontransformed CD4 $^{+}$ T $_{H}$ 1 T-cell clone that is stimulated by pigeon cytochrome *c* peptide (amino acids 81–104) in the context of E k (refs 11, 14). A.E7 cells (2×10^4) pre-treated for 48 h with 100 units per ml of IL-2 and washed 3 times with medium (Click's medium with 10% fetal calf serum, 2 mM glutamine and 50 μ M β -mercaptoethanol; Biofluids) were

stimulated in 200 μ L with 1×10^4 DCEK (E k , E k -expressing L-cell transfectants; gift from R. Germain) antigen-presenting cells given 3,000 rads and 1 μ M peptide. For photomicroscopy, cells were stained after 40 h stimulation with 0.4% trypan blue in phosphate-buffered saline (PBS: 0.8 mM potassium phosphate, 154 mM NaCl and 2.9 mM sodium phosphate, pH 7.4) for 10 min and washed 3 times with PBS. Photomicrographs were taken on a Zeiss Axiovert 405 M microscope using Hoffman optics. Antibody stimulations were done in wells coated with 10 μ g ml $^{-1}$ solutions of protein A-column-purified anti-CD ϵ antibody (145-2C11; gift from P. Kehn) 15 . Each well contained 5×10^4 A.E7 cells in 200 μ L. Lymph node cells from BALB/c mice were treated with 3 μ g ml $^{-1}$ concanavalin A for 48 h, washed with 10 mg ml $^{-1}$ α -methylmannoside, and cultured for 48 h with either 3 or 100 units of IL-2. 1×10^7 cells in 12 ml medium containing either 3 or 100 units IL-2 per ml were stimulated in culture flasks coated with antibodies F23.1 and RR4-7, as described in Table 2 legend. DNA preparations from A.E7 cells stimulated with IL-2 and anti-CD3 ϵ have been described 5 .

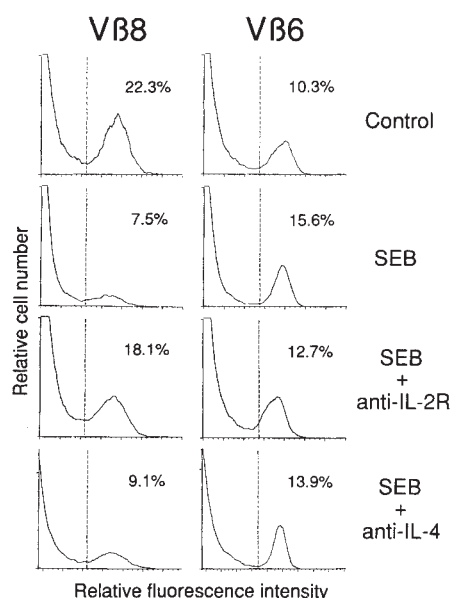


FIG. 2 IL-2 dependent clonal elimination of V β 8 T cells by immunization with SEB. Flow cytometry histograms of lymph-node T cells from uninjected mice (control), or mice injected with SEB (SEB), or with SEB plus antibody 3C7, which blocks the binding of IL-2 to the IL-2 receptor α chain (SEB + anti-IL-2R), or with SEB plus antibody 11B11, which blocks IL-4 (SEB + anti-IL-4). V β 8 samples (left panels) were stained with antibody KJ16-133 (ref. 20), which detects V β 8.1, 8.2 receptors, and V β 6 samples (right panels) were stained with monoclonal antibody RR4-7 (ref. 18). Histograms show relative fluorescent intensity versus cell number; positive cells, gated by the dotted line, are percentages of total T cells.

METHODS. Female 6-week-old BALB/c mice were injected with *Staphylococcus* enterotoxin B (Sigma) in 250 μ L PBS in the tail vein as follows: day 0, 500 μ g; day 2, 125 μ g; and day 4, 125 μ g. Intraperitoneal injections of 800 μ L 3C7 were initiated 12 h after the first SEB dose and given every 12 h for 8 days. Antibody 11B11 was given likewise in 1-mg doses. Antibodies were injected in 300 μ L 5% dextrose in water, which provided a simple metabolite and hydration to prevent mortality. After eight days, lymph node cells were stained in 1 $\times$ PBS with 0.1% bovine serum albumin and 0.1% sodium azide with pre-determined dilutions of primary antibody (mouse immunoglobulin Ig γ 2a antibody F23.1 or rat IgG2b, KJ16-133 for V β 8 or rat IgG2b RR4-7 for V β 6), followed by either a fluorescein isothiocyanate-conjugated goat anti-mouse IgG2a (Southern Biotechnology) or a goat F(ab') $_2$ anti-rat IgG H and L (Caltag). CD4 was detected with phycoerythrin-conjugated anti-L3T4 antibody (Becton-Dickinson). Minor residual dead cells were gated by propidium iodide. 50,000 events were analysed on a Becton-Dickinson FACS 440 dual laser cytometer interfaced to a Digital Equipment Corporation PDP 11/24 computer.

larger amounts of SEB are repeatedly administered than with a single dose (data not shown). This supports the model that antigen re-stimulation of T cells under the influence of IL-2 will cause apoptosis. The responding T cells are not completely ablated, suggesting that some might escape death and become memory cells. How does IL-2, which is well known as a T-cell mitogen, paradoxically program the same cell type for apoptosis? One possibility is that IL-2 serves only to drive T cells into the division cycle that may predispose thymocytes and $\gamma\delta$ T cells to death^{27,28}. If so, then any successful immune response could predispose mature $\alpha\beta$ T cells to apoptosis. Alternatively, IL-2 could provide a qualitatively or quantitatively distinct signal that entrains apoptosis to antigen-receptor stimulation. It will be important in either case, in evaluating IL-2 as a therapeutic agent in humans, to consider its unexpected ability to predispose T cells to apoptosis. □

Received 13 May; accepted 11 September 1991.

1. Kawabe, Y. & Ochi, A. *Nature* **349**, 245–248 (1991).
2. Russell, J. H., White, C. L., Loh, D. Y. & Meleedy-Rey, P. *Proc. natn. Acad. Sci. U.S.A.* **88**, 2151–2155 (1991).
3. Liu, Y. & Janeway, C. A. Jr *J. exp. Med.* **172**, 1735–1739 (1990).
4. Webb, S., Morris, C. & Sprent, J. *Cell* **63**, 1249–1256 (1990).

5. Smith, C. A., Williams, G. T., Kingston, R., Jenkinson, E. R. & Owen, J. T. *Nature* **337**, 181–184 (1989).
6. MacDonald, H. R. & Lees, R. K. *Nature* **343**, 642–644 (1990).
7. Murphy, K. M., Heimberger, A. B. & Loh, D. Y. *Science* **250**, 1720–1723 (1990).
8. Jones, L. A., Chin, L. T., Longo, D. L. & Kruisbeek, A. M. *Science* **250**, 1726–1729 (1990).
9. Rocha, B. & von Boehmer, H. *Science* **251**, 1225–1228 (1991).
10. Hecht, T. T., Longo, D. L. & Matis, L. A. *J. Immun.* **131**, 1049 (1983).
11. Ashwell, J. D., Robb, R. J. & Malek, T. R. *J. Immun.* **137**, 2572–2578 (1986).
12. Nau, G. J., Moldwin, R. L., Landoli, D. W., Kim, D.-K., & Fitch, F. W. *J. Immun.* **139**, 114–122 (1987).
13. Williams, M. E., Lichtman, A. H. & Abbas, A. K. *J. Immun.* **144**, 1208–1214 (1990).
14. Weaver, C. T., Hawrylycz, C. M. & Unanue, E. R. *Proc. natn. Acad. Sci. U.S.A.* **85**, 8181–8185 (1988).
15. Mueller, D. L., Jenkins, M. K. & Schwartz, R. H. *J. Immun.* **142**, 2617–2628 (1989).
16. Leo, O., Foo, M., Sachs, D. H., Samelson, L. E. & Bluestone, J. A. *Proc. natn. Acad. Sci. U.S.A.* **84**, 1374–1378 (1987).
17. Staerz, U. D., Rammensee, H.-G., Benedetto, J. D. & Bevan, M. J. *J. Immun.* **134**, 3994–4000 (1985).
18. Kanagawa, D. *J. exp. Med.* **170**, 1513–1519 (1989).
19. White, J. *et al. Cell* **56**, 27–35 (1989).
20. Haskins, K. *et al. J. exp. Med.* **160**, 452–471 (1984).
21. Malek, T. R., Ortega, R. G., Jakway, J. P., Chan, C. & Shevach, E. M. *J. Immun.* **133**, 1976–1982 (1984).
22. Tentori, L., Longo, D. L., Zühiga-Pflücker, J. C., Wing, C. & Kruisbeek, A. M. *J. exp. Med.* **168**, 1741–1747 (1988).
23. Ohara, J. & Paul, W. E. *Nature* **315**, 333–336 (1985).
24. Finkelman, F. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **83**, 9675–9678 (1986).
25. Monod, J. & Jacob, F. *Cold Spring Harbor Symp. quant. Biol.* **26**, 389–401 (1961).
26. Marrack, P., Blackman, M., Kushner, E. & Kappler, J. *J. exp. Med.* **171**, 455–464 (1990).
27. Penit, C. & Vasseur, F. *J. Immun.* **140**, 3315–3323 (1988).
28. Janssen, O. *et al. J. Immun.* **146**, 35–39 (1991).
29. Crispe, I. N., Bevan, M. J. & Staerz, U. D. *Nature* **317**, 627–629 (1985).

ACKNOWLEDGEMENTS. We thank R. Germain, M. Grilli, A. Kuang, A. Kruisbeek, K. Lee, D. Margulies, W. Paul, E. Shevach, L. Staudt and J. Zühiga-Pflücker for advice and assistance, S. Kang and M. Rosner for comments on the manuscript, and A. Tran, K. Novak, and F. Hausman for technical assistance. M.J.L. is a recipient of an Investigator Award from the Cancer Research Institute.

A distinct *Hox* code for the branchial region of the vertebrate head

Paul Hunt*, Massimo Gulisano†, Martyn Cook*,
Mai-Har Sham*, Antonio Faiella†, David Wilkinson*,
Edoardo Boncinelli† & Robb Krumlauf*‡

* MRC Laboratory of Eukaryotic Molecular Genetics, NIMR, The Ridgeway, Mill Hill, London NW7 1AA, UK

† IIGB, CNR, Via Marconi 12, 80125 Naples, Italy

THE branchial region of the vertebrate head forms through complex interactions involving rhombomeric segments, neural crest and branchial arches¹. It is thought that aspects of their patterning mechanisms are linked² and involve *Hox-2* genes, whose overlapping and spatially restricted expression domains represent a combinatorial code for generating regional diversity^{3–5}. Vertebrates possess four *Hox* clusters of *Antennapedia* class homeobox genes, related to each other by duplication and divergence from a common ancestral complex^{3,6–8}. In consequence, at equivalent positions in different clusters there are highly related genes known as subfamilies or paralogous groups. As *Hox-2* genes cannot fully account for patterning individual rhombomeres, we investigated whether offsets in expression limits of paralogous genes could account for the generation of regional diversity. We report here that, with the exception of the *labial* subfamily, paralogues show identical expression limits in rhombomeres, cranial ganglia and branchial arches, providing a combinatorial *Hox* code for the branchial region that seems to be different in organization to that of the trunk.

We isolated and identified two new genes, *Hox-1.11* and *Hox-4.9* for this study. Extensive screening both in mice and in humans suggests that there are only 12 members in the four 3' *Hox* subfamilies, and we have analysed the expression of 11 of these in mouse embryos. All genes in a given subfamily, with the exception of *labial*, are expressed with the same rhombomere boundary (Fig. 1). The *Dfd* subfamily map to r6/r7 (Fig. 1a–c), the *Zen/pb* genes to r4/r5 (Fig. 1d–f), and the *pb* subfamily to

r2/r3 (Fig. 1g, h). The boundaries between genes in *Hox-1* and *Hox-4* also vary with a two-segment periodicity, and there is no evidence for an offset in expression limits between members of a subfamily. But subfamily members can show distinct differences within their domains of expression. For example, *Hox-1.5* and *Hox-2.7* both show identical patterns of expression, with higher levels in r5 and lower levels posteriorly (Fig. 1d, e), whereas *Hox-4.1* shows a pattern complementary to this in the same overall domains (Fig. 1f). At this stage all members of the *labial* subfamily show significant differences between each other both in sites and in spatial domains of expression. At 9.5 days of development only *Hox-2.9* (refs 4, 9, 10) is expressed in the hindbrain (Fig. 1l). An equivalent section hybridized with *Hox-1.6* (Fig. 1j) shows no expression in the rhombomeres and cranial ganglia. *Hox-4.9* also shows no specific localization in the neural tube (Fig. 1k), but is expressed in patches of surface ectoderm overlying the hindbrain.

We have extended these comparisons to other tissues in the head, and show that members of a subfamily also have identical spatial restrictions in neural crest derivatives and branchial arches. This is illustrated for two subfamilies in Fig. 2. In the *Pb* subfamily, *Hox-1.11* and *Hox-2.8* hybridize to the vii/viii ganglion complex (Figs 2f and 1g, h) and show expression in the mesenchyme of the second and third branchial arches (Fig. 2e, d). In the *Zen/Pb* subfamily, *Hox-1.5* (Figs 1e and 2b, c) and *Hox-2.7* (Fig. 2a) show expression in the third and fourth branchial arches, their overlying surface ectoderm and the ix/x ganglion complex. Previous work and our own data suggest that there is only limited expression of the *labial* genes in neural crest derivatives^{5,10–12}.

The patterns of expression in rhombomeres and branchial arches at 9.5 days are a result of modulations of patterns established at earlier stages of development. In a one-somite embryo, *Hox-2.8* has a restricted domain of expression that respects the presumptive r2/r3 boundary (Fig. 3b), as defined by the early expression domain of *Krox 20* shown in a near adjacent section (Fig. 3c). It also seems to have a similar expression domain in a presomitic embryo (Fig. 3a). This early boundary is identical to that at later stages, and shows that *Hox-2.8* expression is established before overt rhombomere formation in presomite embryos. Previous work has shown that members of the *Zen/Pb* and *Dfd* subfamilies of *Hox-2* are not at their final anterior limits at this stage, but have been established by 8.5 days of development (8–12 somites)⁴. We think that the differences

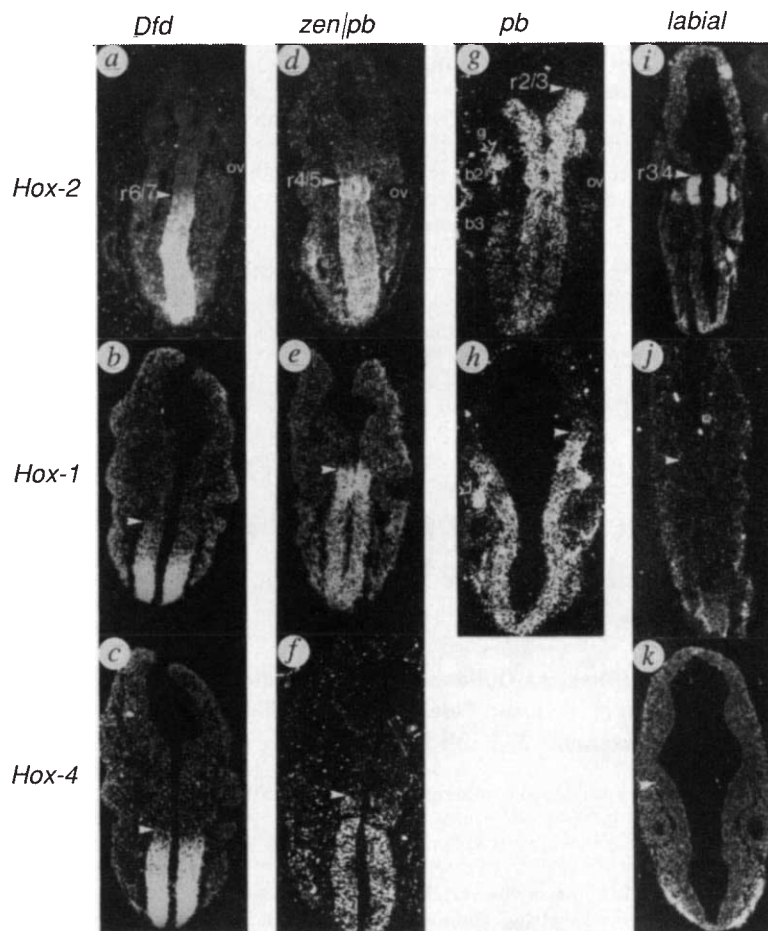
‡ To whom correspondence should be addressed.

between these three subfamilies, summarized in Fig. 4a, reflect a temporal progression in the establishment of rhombomeric expression boundaries.

The *labial* subfamily is also different in that the later patterns of expression (described above) are dramatically different from earlier stages (Fig. 4a). At 8 days of development *Hox-1.6* and *Hox-2.9* are expressed in a continuous domain up to the pre-

sumptive r3/r4 boundary¹¹. But at a similar stage of development (the same one-somite embryo as Fig. 3b, c) *Hox-4.9*, the third member of the family, is only expressed in the most posterior regions of the embryo (Fig. 3d). In contrast to other subfamilies, these early *labial* patterns are not maintained; in later stages only *Hox-2.9* shows expression in the hindbrain and a limited subset of cranial neural crest^{3,10} (Figs 1i-k and 4a, b).

FIG. 1 Expression of genes of the *Dfd*, *zen/pb*, *pb* and *labial* subfamilies in the 9.5-day mouse embryo hindbrain. The *Dfd* group consists of *Hox-2.6* (a), *Hox-1.4* (b), *Hox-4.2* (c) and *Hox-3.5* (not shown). The *Zen/Pb* group consists of *Hox-2.7* (d), *Hox-1.5* (e) and *Hox-4.1* (f). The *Pb* group consists of only two members, *Hox-2.8* (g) and *Hox-1.11* (h). The *labial* group consists of *Hox-2.9* (i), *Hox-1.6* (j) and *Hox-4.9* (k). ov, Otic vesicle; r6/7, boundary between rhombomeres 6 and 7; b2 and b3, second and third branchial arches; g, vii/viii ganglion. In a subgroup the white arrowheads indicate the same rhombomere boundary on each section. The vii/viii ganglion complex is indicated by a white open arrow. Members of a subfamily are expressed in the same rhombomeres, and expression limits of subfamilies differ from each other by two rhombomeres, with the exception of *labial*. There are differences between subfamily members in amounts of expression in rhombomeres. *Hox-1.4* is expressed in posterior regions of r7 (b), but expression decreases to background amounts at the r6/r7 boundary. *Hox-2.6* is expressed up to this boundary (a), but expression is much weaker in the anterior portion of the rhombomere. *Hox-4.2* is expressed with a smaller decrease at the r6/r7 boundary (c). *Hox-4.1* has highest amounts of expression in rhombomeres posterior to r6, and a lower amount of expression distinct from background in r5 (f). *Hox-2.7* (d) and *Hox-1.5* (e) show complementary amounts of expression in the same rhombomeres as *Hox-4.1*. *Hox-2.8* has highest amounts of expression in r3, r4 and r5 (g), whereas *Hox-1.11* has highest amounts in r3 and r5 (h). *In situ* hybridization was as described³⁵. Eight probes were as previously described: *Hox-2.6*, 2.7, 2.8 and 2.9 (ref. 4); *Hox-1.4* and -2.4 (refs 13, 36); *Hox-1.5* (ref. 27); *Hox-1.6* (ref. 37). *Hox-4.1* was isolated by polymerase chain reaction according to the protocol of Frohman and Martin³⁸, using oligonucleotide sequences derived from the EMBO database. In mouse only one gene related to *pb* is known, *Hox-2.8*. In human, however, there is also an equivalent in *HOX-1*, *HOX-1K* (ref. 39). Sequence comparison revealed little DNA sequence homology between *Hox-2.8* and *HOX-1K* (data not shown)³⁹, and northern analysis revealed specific hybridization of *HOX-1K* probe to mouse RNA, at a size distinct from *Hox-2.8* transcripts (data not shown). Human probes are able to hybridize specifically to their mouse equivalents under high-stringency *in situ* conditions, and *HOX-1K* probes recognize a distinct band in mouse genomic DNA whose size is different from that containing *Hox-2.8*. A *Hox-4.9* complementary DNA was isolated from an 8.5-day mouse embryo library, using a probe



from the human *HOX-4G* Gene³⁹. Sequence was determined enzymatically using dideoxy nucleotides, and shows greater similarity to human *HOX-4G* than to other *labial* family genes. The peptide sequence of the homeobox is: PSAIRTNFSTKQLTELEKEFHFNKYLTRARRIEIANCLQNDTQVKIWFQNRMRMKQKKRER (single-letter amino-acid code).

FIG. 2 Expression of genes of the *Dfd*, *zen-like* and *pb* subfamilies in the branchial arches of a 9.5-day mouse embryo. a, *Hox-2.7*; b, c, *Hox-1.5*; d, *Hox-2.8*; e, f, *Hox-1.11*. *Hox-2.7* and *Hox-1.5* are expressed in the mesenchyme (mes) of the third and posterior branchial arches, and not in b1 and b2. c, Detail of b, showing that *Hox-1.5* has the same restrictions of expression in branchial arch ectoderm (ect) as *Hox-2.7*. *Hox-2.8* (d) and *Hox-1.11* (e) are expressed in the second and posterior branchial arches, but not in the first. f, Detail of e, showing *1K* expression in the vii/viii ganglion complex and the expression in the neural tube. In contrast to *Hox-2.8*, which shows strong expression in rhombomeres 3, 4 and 5, *1K* shows strongest expression in r3 and r5, with similar levels in r4 to more posterior parts of the hindbrain. Probes as in Fig. 1.

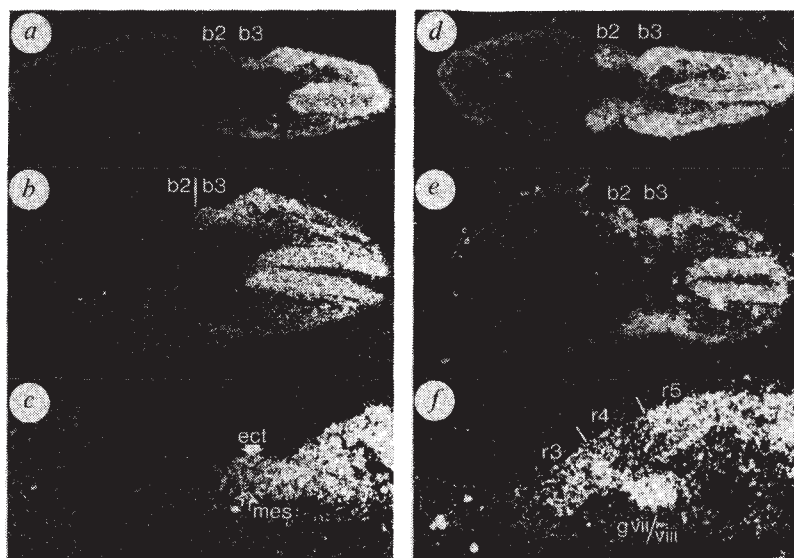
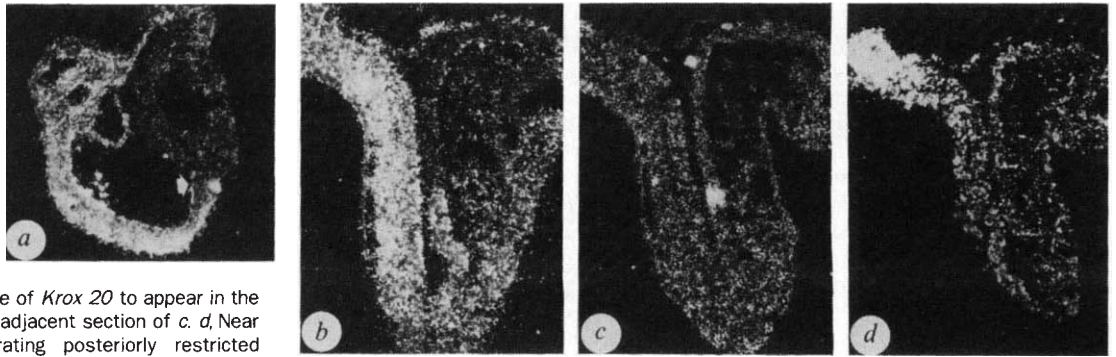


FIG. 3 Expression of *Hox* genes in 8-day mouse embryos. *a, b*, *Hox-2.8*; *c*, *Krox 20*; *d*, *Hox-4.9*. *a*, Sagittal section of a 0-somite embryo, with an anterior limit indicated by a white arrow. This shows a similar localization to the one-somite embryo shown in *b*, where the anterior expression limit corresponds to the first stripe of *Krox 20* to appear in the hindbrain, shown in the near adjacent section of *c*. *d*, Near adjacent section demonstrating posteriorly restricted expression of *Hox-4.9*. Probes as for Fig. 1.

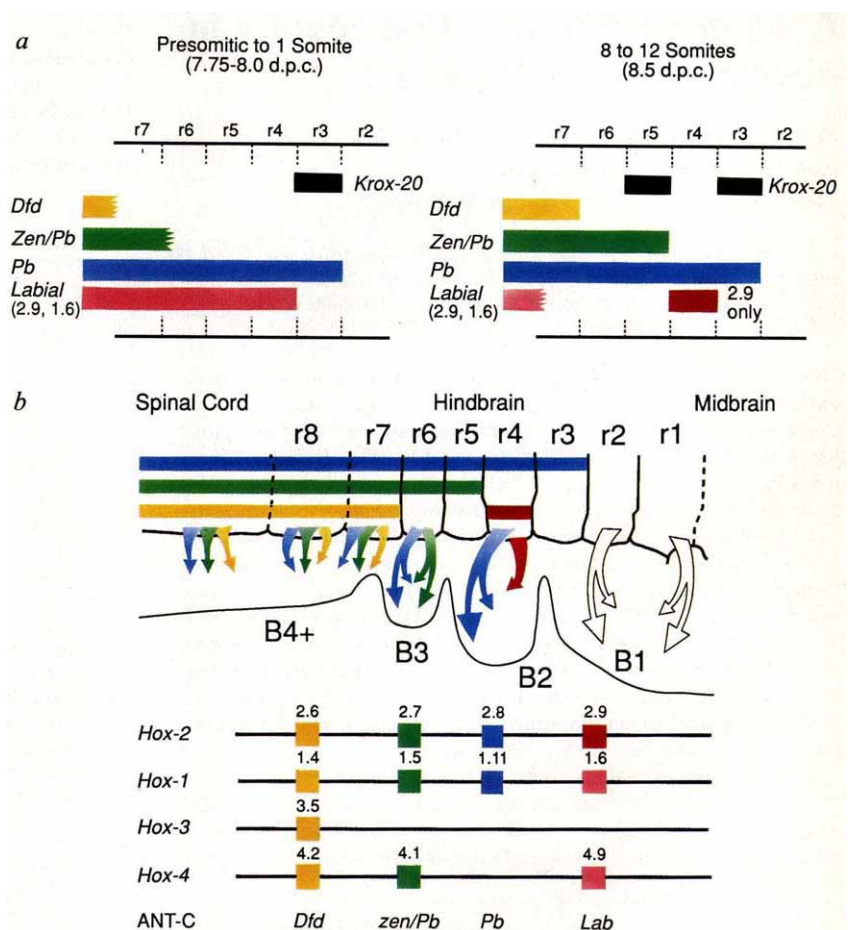


Our results show that, with the exception of *labial*, the *Hox* subfamilies show very similar spatial and temporal expression patterns during critical phases of head morphogenesis (summarized in Fig. 4). This suggests conservation of regulatory elements between *Hox* clusters necessary to generate these expression domains. The high degree of similarity in protein sequence^{13,14} and gene expression in a subgroup¹⁵ could imply early functional compensation or redundancy. On the basis of dorso-ventral distributions¹⁶ and differences between subfamily members in the same domains of expression¹⁵ (Fig. 1), we think that subfamily members are uniquely required for the development of specific tissues. Neural crest migrates from specific rhombomeres^{5,17} and carries a form of pre-patterning, which not only enables it to specify a particular structural pattern, but also enables it to instruct surrounding mesodermal and ectodermal tissue¹. The *Hox* patterns support the idea that by specifying

rhombomere identity, patterning information can be transferred to the branchial arches through neural crest^{5,18,19}. These patterns define a combinatorial *Hox* code for the generation of regional diversity in the branchial region, and make predictions of the effects of altering *Hox* gene expression on mouse development. A combinatorial *Hox* code is also important in the patterning of the limb²⁰⁻²² and paraxial mesoderm of the trunk^{7,23,24}.

Direct support for these predictions comes from mice homozygous for null alleles of *Hox-1.5* (ref. 25) and *Hox-1.6* (ref. 26). *Hox-1.5* mutants show defects in tissues derived from mesenchymal cranial neural crest, yet have normal rhombomeres and cranial ganglia²⁵, despite expression in these tissues (Figs 1*a* and 2*b*)²⁷. By contrast, *Hox-1.6* mutants have defects concentrated in cranial ganglia in the r4/r7 region²⁶. These phenotypes support the suggestion that specific subfamily members are required for patterning particular subsets of cranial neural crest.

FIG. 4 Summary of the establishment of the *Hox* code in the head and its transmission to the branchial apparatus by migrating neural crest. The genes that correspond to the subfamily expression domains indicated are shown at the base of *b* in the same colours. There is evidence from characterization of human *HOX* clusters that there are no *Hox* genes 3' of *Hox-3.5* (ref. 39). The early domain of *Hox-1.6* and *Hox-2.9* is indicated in pink, and the later domain of *Hox-2.9* is indicated in red. We have no evidence for *Hox-4.9* expression in the hindbrain at any stage. *a*, Establishment of expression before the appearance of morphologically identifiable rhombomeres. The presumptive boundaries are indicated with dotted lines. At 8 days the *pb* family of genes and *Hox-1.6* and *Hox-2.9* of the *labial* family¹¹ have reached anterior expression boundaries that show the same relationship to the rhombomere 3 *Krox 20* stripe as they do when overt rhombomere morphology appears. At this stage genes of the *Dfd* and *zen/pb* groups have not yet reached their anterior limits⁴ indicated by the broken anterior limits, consistent with the idea of temporal collinearity of establishing expression with position of a gene in a cluster. At 8.5 days when two *Krox 20* stripes are emerging, the differences in expression between *Hox-2.9* and *Hox-1.6* are apparent¹¹. The posterior domains of both genes are receding caudally, whereas only *Hox-2.9* shows an expression domain restricted to r4. Genes of the other 3' subfamilies are now all expressed with the same relationships to each other and *Krox 20* as at later stages. At this stage neural crest migration is occurring. *b*, Expression pattern after neural crest migration is complete, when distinct rhombomeres are apparent. *Krox 20* is expressed at this stage, but its expression domain is omitted for clarity. The migration of neural crest, indicated by coloured arrows, has transferred a combinatorial code of *Hox* subfamily expression to the cranial ganglia and branchial arches. The short red arrow indicates that *Hox-2.9* expression is confined to the ganglionic crest.



In the *Hox-1.5* and *-1.6* mutants there was no evidence for transformations of structures, in contrast to the transformations of cervical vertebrae seen when *Hox-1.1* is ectopically expressed in somites²⁴. Because the development of the head involves interactions between rhombomeres, neural crest, surface ectoderm, paraxial mesoderm and pharyngeal endoderm, the alterations of *Hox* expression in a subset of these might not generate a simple transformation, particularly as the role of *Hox* genes in patterning the mesoderm and endoderm is not clear. Furthermore, removal of one *Hox* gene may alter expression of multiple *Hox* genes and other components of the head specification network.

The *Hox* code of the branchial region is different from that of the trunk, where anterior boundaries of subfamily members can be offset from each other. There are many morphological grounds for believing that the head and trunk have distinct developmental mechanisms^{1,18,28,29}, which we believe has resulted in the use of the same genes in different ways in the two contexts. *Antennapedia* class *Hox* genes are not expressed in more anterior parts of the head, which must therefore employ other genes³⁰⁻³² and patterning mechanisms²⁸. In an analogous way the patterning of anterior parts of the head in *Drosophila* is also thought to involve molecular mechanisms independent of *Antennapedia* class genes^{33,34}. □

Received 1 August; accepted 26 September 1991.

1. Noden, D. *Development* **103** (suppl.), 121-140 (1988).
2. Lumsden, A. *Trends Neurosci.* **13**, 329-335 (1990).
3. Graham, A., Papalopulu, N. & Krumlauf, R. *Cell* **57**, 367-378 (1989).
4. Wilkinson, D. *et al. Nature* **341**, 405-409 (1989).
5. Hunt, P., Wilkinson, D. & Krumlauf, R. *Development* **112**, 43-51 (1991).
6. Duboule, D. & Dolle, P. *EMBO J.* **8**, 1497-1505 (1989).

7. Kessel, M. & Gruss, P. *Science* **249**, 374-379 (1990).
8. Boncinelli, E. *et al. Trends Genet.* **7**, 329-334 (1991).
9. Murphy, P., Davidson, D. & Hill, R. *Nature* **341**, 156-159 (1989).
10. Frohman, M., Boyle, M. & Martin, G. *Development* **110**, 589-607 (1990).
11. Murphy, P. & Hill, R. *Development* **111**, 61-74 (1991).
12. Sundin, O. & Eichele, G. *Genes Dev.* **4**, 1267-1276 (1990).
13. Featherstone, M. S. *et al. Proc. natn. Acad. Sci. U.S.A.* **85**, 4760-4764 (1988).
14. Graham, A. *et al. Genes Dev.* **2**, 1424-1438 (1988).
15. Gaunt, S. J., Krumlauf, R. & Duboule, D. *Development* **107**, 131-141 (1989).
16. Graham, A., Maden, M. & Krumlauf, R. *Development* **112**, 255-264 (1991).
17. Lumsden, A., Sprawson, N. & Graham, A. *Development* (in the press).
18. Lumsden, A. in *Seminars in Developmental Biology, The Evolution of Segmental Patterns* Vol. 1 (ed. Stern, C.) 117-125 (Saunders, Philadelphia, 1990).
19. Hunt, P. *et al. Development* **112**, (suppl.) 187-196 (1991).
20. Dolle, P. *et al. Nature* **342**, 767-772 (1989).
21. Izpisua-Belmonte, J.-C. *et al. Nature* **350**, 585-589 (1991).
22. Nohno, T. *et al. Cell* **64**, 1197-1205 (1991).
23. Kessel, M. & Gruss, P. *Cell* **67**, 89-104 (1991).
24. Kessel, M., Balling, R. & Gruss, P. *Cell* **61**, 301-308 (1990).
25. Chisaka, O. & Capecchi, M. *Nature* **350**, 473-479 (1991).
26. Lufkin, T. *et al. Cell* **66**, 1105-1119 (1991).
27. Gaunt, S. J. *Development* **101**, 51-60 (1987).
28. Thorogood, P. *Development* **103**, 141-153 (1988).
29. Holland, P. in *Seminars in Developmental Biology, The Evolution of Segmental Patterns* Vol. 1 (ed. Stern, C.) 135-145 (Saunders, Philadelphia, 1990).
30. McMahon, A. & Bradley, A. *Cell* **62**, 1073-1085 (1990).
31. Dolle, P. *et al. Development* **110**, 1133-1151 (1990).
32. Price, M. *et al. Nature* **351**, 748-751 (1991).
33. Cohen, S. & Jürgens, G. *Nature* **346**, 482-485 (1990).
34. Finkelstein, R. & Perrimon, N. *Nature* **346**, 485-488 (1990).
35. Wilkinson, D. & Green, J. in *Postimplantation Mouse Embryos: A Practical Approach* (eds Rickwood, D. & Cockcroft, D. L.) 155-171 (IRL, Oxford, 1990).
36. Gaunt, S. J. *Development* **103**, 135-144 (1988).
37. Duboule, D. *et al. EMBO J.* **5**, 1973-1980 (1986).
38. Frohman, M. & Martin, G. *Technique* **1**, 165-170 (1989).
39. Simeone, A. *et al. Mech. Dev.* **33**, 215-227 (1991).

ACKNOWLEDGEMENTS. We would like to thank T. Lufkin and P. Chambon for sharing data before publication on the *Hox 1.6* mutants, D. Duboule for the gift of mouse *Hox 1* probes, P. Sharpe for information on the lack of paralogues in the mouse *Hox 3* complex, D. Noden, P. Thorogood and A. Lumsden for discussions on head development, members of the Boncinelli lab for sharing information on human *HOX* clusters, L. Stubbs for chromosomal localization of *Hox 4.9* and J. Brock for graphic illustration. This work was funded in part by the UKMRC, the 3rd AIDS project and the Italian Association for Cancer Research. P.H. is in receipt of an MRC studentship.

A de novo Alu insertion results in neurofibromatosis type 1

Margaret R. Wallace*, Lone B. Andersen, Ann M. Saulino, Paula E. Gregory†, Thomas W. Glover† & Francis S. Collins‡

Howard Hughes Medical Institute, and the Departments of Internal Medicine and Human Genetics, and †Pediatrics and Communicable Diseases, University of Michigan, 4570 MSRB II, 1150 W. Medical Center Drive, Ann Arbor, Michigan 48109-0650, USA

NEUROFIBROMATOSIS type 1 (NF1) is a common autosomal dominant disorder with a high mutation rate and variable expression, characterized by neurofibromas, café-au-lait spots, Lisch nodules of the iris, and less frequent features including bone deformities and learning disabilities¹. The recently cloned *NF1* gene encodes a transcript of 13 kilobases from a ubiquitously expressed locus on chromosome 17 (refs. 2-4). Most NF1 patients are expected to have unique mutations, but only a few have so far been characterized, restricting genetic and functional information and the design of DNA diagnostics. We report an unusual *NF1* mutation, that of a *de novo* Alu repetitive element insertion into an intron, which results in deletion of the downstream exon during splicing and consequently shifts the reading frame. This previously undescribed mechanism of mutation indicates that Alu retrotransposition is an ongoing process in the human germ line.

The 31-year-old male patient (D.D.) exhibits several features of NF1, including one cutaneous neurofibroma, axillary freckling, Lisch nodules, cervical nerve root tumours, and macrocephaly. Café-au-lait spots are not present. His parents

show no signs of NF1, and DNA fingerprinting analysis found no evidence of nonpaternity. Part of the *NF1* complementary DNA detected an abnormal Southern blot pattern in the patient's DNA after digestion with several restriction enzymes². This was consistent with a small insertion (300-500 basepairs (bp)) in a 3.8-kilobase (kb) *EcoRI* fragment which contains six *NF1* exons

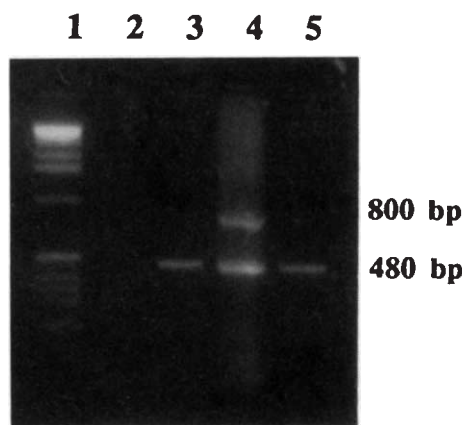


FIG. 1 Ethidium bromide staining of a 1.0% agarose gel demonstrates the insertion in the exon 6 PCR product. Lane 1 contains the BRL 1-kb ladder, lane 2 contains a water (negative) control, lanes 3 and 5 are products from the patient's father and mother, respectively, and the patient's PCR products are shown in lane 4. All show the normal fragment of 480 bp, but the patient also has an abnormal fragment of ~800 bp. DNA from both the patient's leukocytes and from an established lymphoblastoid line gave the same result (data not shown).

METHODS. Genomic DNA from the patient and his parents was extracted as described². Genomic DNA (100-500 ng) was amplified using the exon 6 primers already described³, with 35 cycles (each cycle entailed 1 min each at 94°C for denaturation, 65°C for annealing, and 72°C for extension) using standard buffers and reagents recommended by Cetus. One-tenth of each PCR reaction was loaded per lane.

* Present address: Center for Mammalian Genetics, Department of Pediatrics, University of Florida Health Science Center, Gainesville, Florida 32610, USA.

‡ To whom correspondence should be addressed.

FIG. 2 Sequence of the *Alu* insertion in the exon-6 PCR product from NF1 patient D.D. The entire sequence of the PCR product was obtained and the exon and other intron sequences were found to be normal, with the exception of the *Alu* insertion. This *Alu* sequence is identical to the PV or HS-1 *Alu* subfamily consensus¹²⁻¹⁴, with the exception of the two bases underlined (a substitution of A for G at position 72, and an additional A in the string of As at position 128). A poly(A) tract over 40 nucleotides long was found at the 3' end.

METHODS. The PCR product was cloned into the *Bam*HI site of Bluescript I (KS-) at cloning sites built into the PCR primers, and transformed into BRL DH5- α cells as recommended by the supplier. Plasmid DNA from two independent clones was extracted¹⁸ and sequenced on both strands by direct sequencing of plasmid DNA (Sequenase, US Biochemicals).

```

1  GGCCGGGCGCGGTGGCTCA•GCCTGTAAT•CCAGCACTTTGGAGGCCGAGCGGGCGGA
61  TCACGAGGTCA•AGAGATCGAGACCATCCGGCTAAACGGTGAAACCCGCTCTACTAA
121 AAATACAAAAA•TTAGCCGGGCGTAGTGGCGGGCGCCTGTAGTCCAGCTACTTGGGAG
181 GCTGAGGCAGGAGAATGGCGTGAACCCGGGAGGCGGAGCTTGCACTGAGCCGAGATCCCG
241 CCACTGCACTCCAGCCTGGGCGACAGAGCGAGACTCCGTCTC(A)n

```

in the 3' third of the *NF1* coding region³. These six exons were individually examined by polymerase chain reaction (PCR) amplification of the patient's DNA, using primers derived from surrounding introns³. One exon, termed number 6 in current nomenclature³, showed a pattern consistent with a 320-bp insertion (Fig. 1, lane 4); the remaining exons were normal. The parents' DNA showed only the normal pattern for exon 6 (Fig. 1, lanes 3 and 5). Analysis of 50 other NF1 patients failed to show this abnormality. The abnormal PCR product was cloned and sequenced; it consisted of the normal exon 6 sequence with an *Alu* repetitive element inserted 44 bp upstream of the exon. Sequencing of this element (Fig. 2) showed that the *Alu* is inserted in the opposite orientation of the *NF1* gene, that a substantial poly(A) tract is present, and that the *Alu* is flanked by 3–13-bp direct repeats (exact length was indeterminate owing to the poly(A) tract) (Fig. 3). The site of insertion was immediately adjacent (upstream) to an A/T stretch of 26 bp, consistent with the proposed sites preferred for *Alu* integration⁵.

As the exonic sequence is undisturbed, we thought that the *Alu* insertion might disrupt normal splicing of the transcript. It was suspected that exon 6 might be lost during splicing, because the insertion was closest to this exon and could interfere with branch-point recognition. A PCR experiment was designed to examine the patient's *NF1* RNA from exons 5 through to 7 (Fig. 3). The normal product, from the beginning of exon 5 through to the middle of exon 7, was expected to be 515 bp long, whereas a corresponding product lacking exon 6 would be 235 bp. RNA PCR consistently gave products of both lengths in the patient's RNA; normal controls showed only the normal band (Fig. 4a). Both products hybridized with exon 5 (Fig. 4b), and exon 6 hybridized with only the normal 515-bp fragment

(Fig. 4c). The 235-bp product was cloned and sequenced and found to contain the sequences expected from exons 5 and 7 precisely spliced together (data not shown). This results in a shift in the translational frame when the exon 7 sequence is encountered, with a predicted subsequent premature truncation of the NF1 protein. Thus, the patient's *NF1* allele containing the *Alu* insertion should produce an NF1 protein that lacks the portion encoded by exon 6 and terminates within the exon 7-encoded region. This would result in a protein missing the C-terminal 771 amino acids of the predicted 2,818 for the entire NF1 protein⁶.

Two somatic cell hybrids containing only the mutant *NF1* allele were constructed on a hamster cell background. RNA PCR detected the abnormal band in the hybrids, with no evidence for production of the normal allele from the mutant gene by hybridization (data not shown). In addition, Southern blot analysis of hybrid and parental DNA using the polymorphic probe THH59 (17q23–25.3) showed that the allele receiving the *Alu* insertion was paternal in origin (data not shown). This agrees with a previous genetic analysis which indicates a pre-dominance of paternal origin in new-mutation NF1 cases⁷.

To our knowledge, this is the first report of a disease-causing mutation consisting of a *de novo Alu* insertion: *Alu* elements

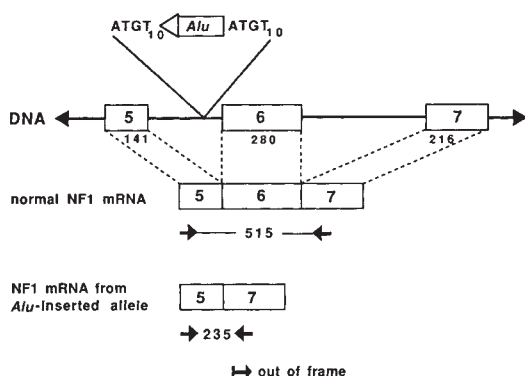


FIG. 3 Schematic of the *Alu* insertion and experimental design for RNA PCR analysis. The *Alu* resides in the intron in the opposite orientation from the *NF1* gene, flanked by direct repeats, 44 bp upstream of exon 6. PCR primers in exons 5 and 7 were chosen to analyse whether exon 6 is present in all of the patient's *NF1* transcripts, or whether the *Alu* insertion causes exon 6 to be lost during splicing of the mutant allele.

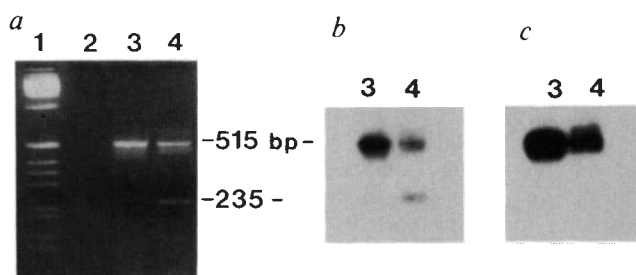


FIG. 4 a, Ethidium bromide staining of PCR products from RNA PCR analysis. Lane 1 contains the BRL 1-kb ladder; lane 2, a water control; lane 3, a normal lymphoblastoid control; and lane 4 shows the products from the patient's analysis. The normal product is 515 bp, and the abnormally small band in the patient's lane is 235 bp. (Other non-NF1 RNA sources analysed showed only the normal fragment (data not shown).) Hybridization with exon 5 (b) indicates that the 235-bp fragment is a product from the *NF1* gene including exon 5. But the 235-bp fragment fails to hybridize to the exon 6 probe (c), indicating that exon 6 is lost during splicing.

METHODS. RNA extracted from lymphoblastoid cell lines¹⁸ was reverse-transcribed with oligod(T)¹⁹. The PCR consisted of 35 cycles of 94 °C denaturation, 65 °C annealing, and 72 °C extension under conditions recommended by Cetus. The exon-5 primer sequence used was 5'-CAGATCTGCTTGATGTTGTAAG-3', and the exon-7 primer sequence was 5'-TGACAGCAGCTGACTTGACTTTGC-3'. The gel was transferred to two membranes (bidirectionally) with standard Southern blotting methods. The filter in b was hybridized to a radiolabelled exon-5 PCR product, and the other membrane (c) was hybridized to a radiolabelled exon-6 PCR product, using standard random priming, hybridization and wash conditions².

have been involved in the generation of disease mutations by recombination⁸⁻¹⁰ or point mutation¹¹, but not as new element. The mechanism of the splicing effect of the inserted element is unclear, but exon skipping is indicative of a defect in branch-point recognition.

Comparison of this *Alu* sequence with known subsets of the *Alu* family revealed that this particular element most closely matches the subfamily known as PV (refs 12, 13) or HS-1 (ref. 14), with the differences noted in the Fig. 2 legend. This subfamily is considered the most recently inserted *Alu* subgroup, many members of which are polymorphic in the human population¹⁵. Consistent with recent origin is the finding that this subfamily is transcriptionally active¹². The close agreement of the *de novo* *Alu* element sequence reported here with this subfamily strongly supports the concept of one or a few 'master' *Alu* elements capable of providing new retrotransposons^{13,15-17}. The presence of the long poly(A) tract (a presumed remnant from an RNA intermediate) and the direct repeats (which presumably arose during integration), as well as the *de novo* appearance of this *Alu*, indicate that it probably arose by retrotransposition in the father's germline. The possibility of insertion into the paternal chromosome during early embryogenesis must also be considered, though no evidence for mosaicism was encountered. If *Alu* retrotransposition is an ongoing process in human

biology, it is likely that there will be other examples of this mechanism of mutation in human genetic disease. □

Received 19 August; accepted 10 September 1991.

1. Riccardi, V. M. & Eichner, J. E. *Neurofibromatosis: Phenotype, Natural History and Pathogenesis* (Johns Hopkins, Baltimore, 1986).
2. Wallace, M. R. *et al. Science* **249**, 181-189 (1990).
3. Cawthon, R. M. *et al. Cell* **62**, 193-201 (1990).
4. Viskochil, D. *et al. Cell* **62**, 187-192 (1990).
5. Daniels, G. R. & Deininger, P. L. *Nucleic Acids Res.* **13**, 8939-8954 (1985).
6. Marchuk, D. A. *et al. Genomics* (in the press).
7. Jadayel, D. *et al. Nature* **343**, 558-559 (1990).
8. Lehrman, M. A., Goldstein, J. L., Russell, D. W. & Brown, M. S. *Cell* **48**, 827-835 (1987).
9. Markert, M. L., Hutton, J. J., Wiginton, D. A., States, J. C. & Kaufman, R. E. *J. clin. Invest.* **81**, 1323-1327 (1988).
10. Berkvens, T. M., Van Ormondt, H., Gerritsen, E. J. A., Khan, P. M. & Van Der Eb, A. J. *Genomics* **7**, 486-490 (1990).
11. Mitchell, G. A. *et al. Proc. natn. Acad. Sci. U.S.A.* **88**, 815-819 (1991).
12. Matera, A. G., Hellmann, U. & Schmid, C. W. *Molec. cell. Biol.* **10**, 5424-5432 (1990).
13. Matera, A. G., Hellmann, U., Hintz, M. F. & Schmid, C. W. *Nucleic Acids Res.* **18**, 6019-6023 (1990).
14. Batzer, M. A. & Deininger, P. L. *Genomics* **9**, 481-487 (1991).
15. Batzer, M. A. *et al. Nucleic Acids Res.* **19**, 3619-3623 (1991).
16. Labuda, D. & Striker, G. *Nucleic Acids Res.* **17**, 2477-2491 (1989).
17. Jurka, J. & Milosavljevic, A. *J. molec. Evol.* **32**, 105-121 (1991).
18. Sambrook, J., Fritsch, E. F. & Maniatis, T. in *Molecular Cloning, A Laboratory Manual* 2nd edn (Cold Spring Harbor Press, New York, 1989).
19. Ginsburg, D. *et al. Proc. natn. Acad. Sci. U.S.A.* **86**, 3723-3727 (1989).

ACKNOWLEDGEMENTS. We thank A. Brereton, D. Wood and B. Hall for technical assistance, and G. Matera for critically reading the manuscript. This work was supported by the National Neurofibromatosis Foundation and the NIH. F.S.C. is an investigator of the Howard Hughes Medical Institute.

Complementation of transforming domains in E1a/*myc* chimaeras

Robert Ralston

Chiron Corporation, 4560 Horton Street, Emeryville, California 94608, USA

THE *myc* oncogene is functionally similar to adenovirus E1a in its ability to collaborate with activated *ras* oncogenes to transform primary fibroblasts^{1,2}. The transforming functions of E1a and *myc* have been mapped to two distinct regions in each protein^{3,4}. I investigated the functional similarities between E1a and *myc* by constructing E1a/*myc* chimaeras to discover whether the individual transforming domains of E1a could complement individual *myc*-transforming domains. Transformation assays in rat embryo fibroblasts demonstrated that the N-terminal transforming domain of E1a (CR1; ref. 5) could complement the C-terminal transforming domain of *myc* in *cis*, and that the reciprocal chimaera (N-terminal *myc*/C-terminal E1a) was also active. Chimaeras constructed using domains from transformation-defective mutants of either E1a or *myc* were inactive, indicating that both E1a and *myc* domains contribute to function. These experiments suggest that transformation by *myc* and E1a may involve interactions with common substrates.

To test complementation of the v-*myc* C-terminal domain by E1a, I constructed a series of chimaeras (Fig. 1a) and tested their ability to transform rat embryo fibroblasts (REFs) in collaboration with EJ-*ras*. These experiments showed that E1a/v-*myc* chimaeras containing CR1 could fully transform REFs to tumorigenesis in syngeneic rats (Table 1a). This observation was explored further by construction of chimaeras using a series of deletion mutants in E1a. As summarized in Table 1b, the chimaera EMCCΔ37-51, which lacks the portion of CR1 with homology to human papilloma virus-16 (HPV-16) protein E7 (ref. 6), was inactive. Complementation in *trans* (ref. 7) was not observed (Table 1b), although this could have been due to inefficiency in either cotransfection or complementation.

To determine whether a functional *myc* C-terminal domain was required for complementation of CR1, I constructed E1a/c-*myc* chimaeras using the transformation-defective c-*myc* mutant IN-373 (ref. 4). The E1a/IN-373 chimaera was transformation-defective, whereas the E1a/c-*myc* chimaera was competent

TABLE 1 E1a N-terminal domains containing CR1 can complement the *myc* C-terminal domain

Oncogenes	Foci per plate (experiments 1, 2, 3)	Growth in soft agar	Tumours in rats
(a) EJ- <i>ras</i> +13EMXC	60	+	NT
EJ- <i>ras</i> +12EMXC	120	+	NT
EJ-RAS+EMCA	22	+	2/2
EJ- <i>ras</i> +EMCC	16, 11	+	2/2
EJ- <i>ras</i> +EMNC	10	+	2/2
EJ- <i>ras</i> +LTR- <i>myc</i>	27, 22, 42	+	2/2
EJ- <i>ras</i> +E1a-243	148	+	2/2
EJ- <i>ras</i> +E1a-289	140	+	2/2
EJ- <i>ras</i> alone	0, 0, 0	-	NT
(b) EJ- <i>ras</i> +MLV-EMCC	57, 46, 70	NT	NT
EJ- <i>ras</i> +MLV-EMCCΔ37-51	0, 0	NT	NT
EJ- <i>ras</i> +MLV-EMCCΔ54-72	22, 31	NT	NT
EJ- <i>ras</i> +MLV-EMCCΔ73-82	108	NT	NT
EJ- <i>ras</i> +MLV-EMCCΔ58-81	104	NT	NT
EJ- <i>ras</i> +MLV-EMCCΔ37-51	0, 0	NT	NT
+E1aΔ120-139			
EJ- <i>ras</i> alone	0, 0, 0	NT	NT
(c) EJ- <i>ras</i> +SV-LTR- <i>myc</i>	42	+	NT
EJ- <i>ras</i> +SV-LTR-IN373	0	-*	NT
EJ- <i>ras</i> +SV-EhMCC	23	+	NT
EJ- <i>ras</i> +SV-EM(IN373)	0	-	NT
EJ- <i>ras</i> alone	0	-	NT

a, Constructs shown in Fig. 1a were tested for their ability to transform REFs in collaboration with EJ-*ras*. Focus formation was assayed as described⁴. Average number of foci per plate is shown (four plates per experiment). Results of repeated experiments also are shown. Where indicated, transformation was characterized further by picking foci and assaying their ability to grow in soft agar and to produce tumours in syngeneic rats. Soft agar colonies were produced by seeding about 10⁴ cells into 0.35% agar. Transformed cells were cloned from soft agar colonies and used to test tumorigenesis in syngeneic rats by injecting 10⁷ cells subcutaneously. Tumours of 2 cm typically were produced within 1 week. b, EMCC deletion mutants are shown in Fig. 1b. E1aΔ120-139 was constructed by ligation of blunt-ended *Clal* and *FokI* sites on pIN20 (ref. 20). c, To test the contribution of the *myc* C-terminal domain to transformation by EMCC, the chimaera EM(IN373) was constructed using the human c-*myc* mutant IN-373 which contains a 4-serine-codon insertion at position 373 (ref. 4). The E1a/c-*myc* chimaera EhMCC, containing the wild-type human c-*myc* C-terminal domain was used as a control. The expression vector used for these studies contained the SV40 enhancer and origin of replication. Expression of c-*myc* and IN-373 was driven by a Mo-MLV long terminal repeat. EMCC and EM(IN373) were expressed from the E1a promoter. NT, Not tested.

* One small colony observed which could not be established in culture.

(Table 1c). It seemed possible that if the activity of the E1a/*myc* chimaeras arose through functional similarities between E1a and *myc*, then the reciprocal chimaera (N-terminal *myc*/C-terminal E1a) also should be active. Two v-*myc*/E1a recom-

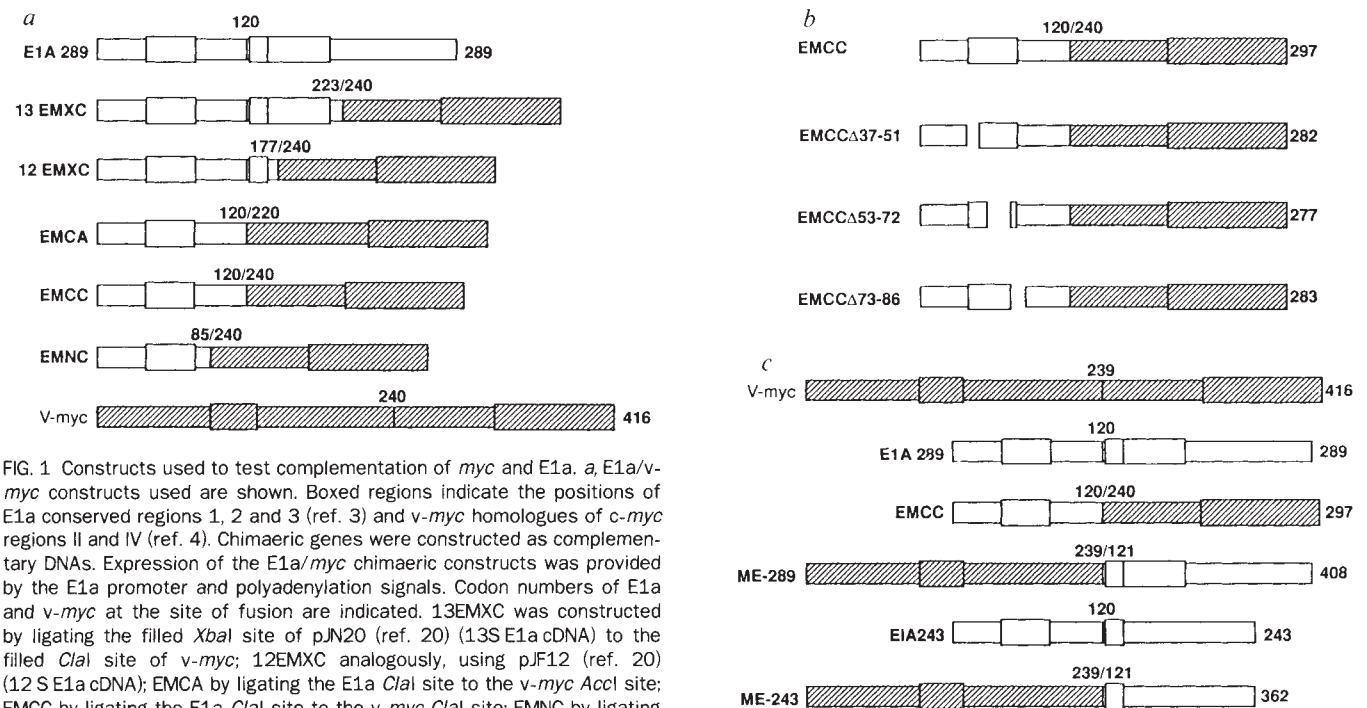


FIG. 1 Constructs used to test complementation of *myc* and E1a. **a**, E1a/*v-myc* constructs used are shown. Boxed regions indicate the positions of E1a conserved regions 1, 2 and 3 (ref. 3) and *v-myc* homologues of *c-myc* regions II and IV (ref. 4). Chimaeric genes were constructed as complementary DNAs. Expression of the E1a/*myc* chimaeric constructs was provided by the E1a promoter and polyadenylation signals. Codon numbers of E1a and *v-myc* at the site of fusion are indicated. 13EMXC was constructed by ligating the filled *Xba*I site of pJN20 (ref. 20) (13S E1a cDNA) to the filled *Clal* site of *v-myc*; 12EMXC analogously, using pJF12 (ref. 20) (12S E1a cDNA); EMCA by ligating the E1a *Clal* site to the *v-myc* *Acc*I site; EMCC by ligating the E1a *Clal* site to the *v-myc* *Clal* site; EMNC by ligating the E1a *Nae*I site to the filled *v-myc* *Clal* site. **b**, Construction of deletion mutants in E1a domain of EMCC. For these and subsequent experiments, the Mo-MLV retroviral vector MX1112 (ref. 21) (gift of M. R. D. Scott) was used to provide expression of the EMCC coding region. Deletion mutants in the E1a N-terminal domain of MLV-EMCC were constructed by overlap-extension polymerase chain reaction (PCR). The E1a mutant domains were subcloned from the amplified DNA and used to replace the intact E1a domain in MLV-EMCC. Accuracy of the mutations was verified by sequence analysis.

The 5' PCR primer used contained a T7 RNA polymerase promoter, allowing the integrity of the coding sequence to be tested also by *in vitro* transcription and translation (data not shown). **c**, Construction of *v-myc*/E1a chimaeras is shown. ME-243 and ME-289 were constructed by ligating the *v-myc* *Clal* site to the *Clal* site of 12S and 13S E1a cDNAs, respectively. MLV-*v-myc* was a gift of A. Bruskin²¹. These constructs were expressed using the vector MX1112.

binants were constructed, using the C-terminal domain of either 13S (289R) or 12S (243R) E1a, and inserted into a Moloney murine leukaemia virus (Mo-MLV) retroviral vector to provide isogenic expression (Fig. 1d).

Both the *v-myc*/E1a chimaera ME-289 and the E1a/*v-myc* chimaera EMCC collaborated with EJ-*ras* to fully transform

REFs (Table 2). By contrast, the *v-myc*/12S E1a chimaera ME-243 was inactive. ME-289-transformed cells produced foci that were less dense than cells transformed by EMCC, E1a or *v-myc*, but cloned efficiently in soft agar and were equally tumorigenic. Also, recombinant retroviruses expressing either ME-289 or EMCC could immortalize primary REFs in the

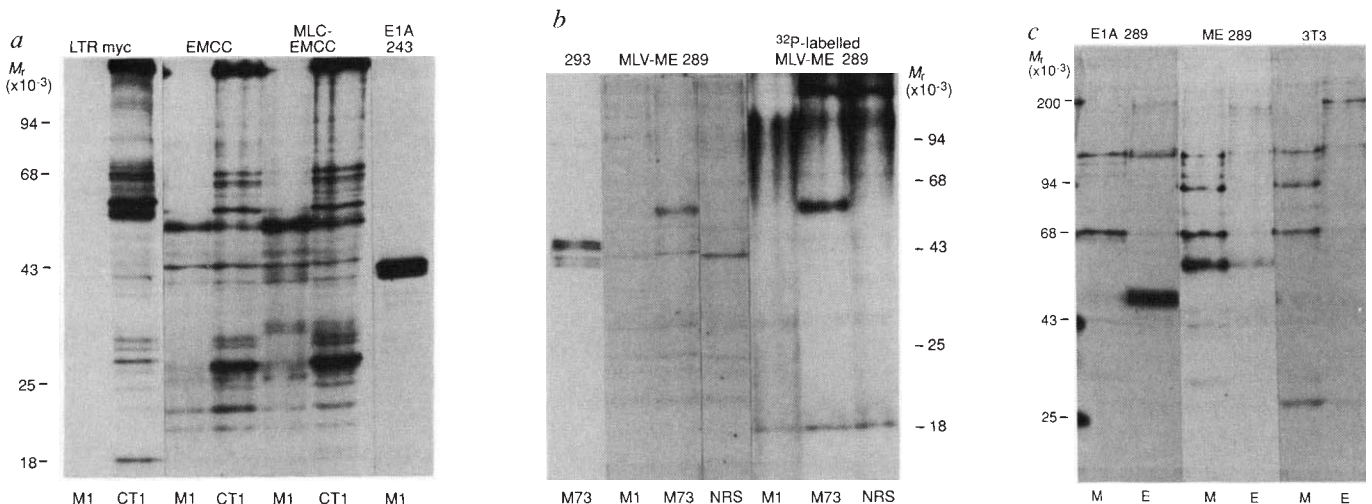


FIG. 2 Immunoprecipitation of EMCC and ME-289 chimaeric proteins. For immunoprecipitation, 5×10^5 cells were labelled for 3 h with 0.5 mCi of 35 S-methionine in 5 ml methionine-free medium. Lysis and immunoprecipitation were done using radioimmunoassay buffer. **a**, Immunoprecipitation of EMCC protein from REFs transformed with EJ-*ras* plus EMCC or MLV-EMCC. Proteins were immunoprecipitated using either rabbit polyclonal antiserum against human *c-myc* C-terminal peptide which crossreacts with *v-myc* (ref. 22) or mouse monoclonal M1, which recognizes the N-terminal domain of E1a (ref. 23). Cells transformed with EM-*ras* plus

either LTR-*myc* or E1a-243 (pJF12) are shown as controls. **b**, Immunoprecipitation of ME-289 protein from REFs transformed with EJ-*ras* plus ME-289. Radiolabelled proteins were precipitated using either mouse monoclonal M1 or M73 (ref. 23), which recognize N-terminal and C-terminal domains of E1a, respectively. Phosphate labelling was used to improve visualization (see ref. 4). **c**, Immunoprecipitation of ME-289 protein from NIH 3T3 cells transformed by recombinant MLV-ME-289 retrovirus, using either rabbit polyclonal antiserum against bacterially-expressed *v-myc* (ref. 24), or mouse monoclonal M73.

TABLE 2 The v-myc N-terminal domain can complement the C-terminal domain of E1a-289

Oncogenes	Foci per plate (experiments 1, 2)	Growth in soft agar	Tumours in rats
Ej-ras + MLV-v-myc	52, 62	+	3/3
Ej-ras + MLV-E1a289	74, 78	+	3/3
Ej-ras + MLV-EMCC	38, 70	+	3/3
Ej-ras + MLV-ME289	12, 29	+	3/3
Ej-ras + MLV-E1a243	165	+	3/3
Ej-ras + MLV-ME243	0, 0-1*	-	NT
Ej-ras alone	0, 0-1*	-	NT

Transformation assays were done as described in Table 1. NT, Not tested.

* Foci did not give rise to established cells.

absence of *ras* (data not shown). Expression of the chimaeric EMCC and ME-289 proteins in transformed cells was demonstrated using antibodies raised against defined regions of E1a and v-myc (Fig. 2). The ability of N-terminal or C-terminal transforming domains of v-myc and Ad5 E1a-289 to complement each other in *cis* suggests that these oncoproteins may interact with common substrates. The molecular basis for common interactions is obscure, however, because mapping studies^{4,8} have not supported a relationship between structurally similar regions⁹⁻¹¹ and transforming activity.

The N-terminal domain of the c-myc protein interacts with Rb the retinoblastoma gene product and HPV-16 E7 can compete with c-myc for Rb binding¹². The requirement of the HPV-16 E7 homology region of CR1 for transformation by E1a/myc chimaeras complements these observations and lends support to the notion that Rb binding is involved in transformation by myc oncogenes. But clear biochemical^{3,13} and biological¹⁴ differences between myc and E1a are apparent from several studies, and even in the case of Rb, myc seems to bind with reduced efficiency relative to E1a (ref. 12). These differences may be reflected in the greater transformation activity of E1a/myc chimaeras compared with myc/E1a chimaeras.

Functional similarities between myc and E1a also may not be apparent under all conditions^{15,16}. For example, c-myc function in normal cells (transcriptional control?) may be regulated by Rb, as suggested for E2F (refs 17, 18); overexpression of c-myc protein in cancer cells may allow myc to titrate Rb, thus providing a novel E1a-like function. Despite these caveats, the apparent interchangeability of transforming domains of E1a and myc suggests that there may be common targets in addition to Rb (ref. 19). The E1a/myc chimaeras described here may provide useful reagents for their identification. □

Received 8 July; accepted 30 August 1991.

- Land, H., Parada, L. F. & Weinberg, R. A. *Nature* **304**, 596-602 (1983).
- Ruley, H. E. *Nature* **304**, 602-606 (1983).
- Moran, E. & Matthews, M. B. *Cell* **48**, 177-178 (1987).
- Stone, J. et al. *Molec. cell. Biol.* **7**, 1697-1709 (1987).
- Kimmelman, D., Miller, J. S., Porter, D. & Roberts, B. E. *J. Virol.* **53**, 399-409 (1985).
- Phelps, W. C., Yee, C. L., Munger, K. & Howley, P. M. *Cell* **53**, 539-547 (1988).
- Moran, B. & Zerler, B. *Molec. cell. Biol.* **8**, 1756-1764 (1988).
- Whyte, P., Ruley, H. E. & Harlow, E. *J. Virol.* **62**, 257-265 (1988).
- Ralston, R. & Bishop, J. M. *Nature* **306**, 803-806 (1983).
- Branton, P. E., Bayley, S. T. & Graham, F. L. *Biochim. biophys. Acta* **780**, 67-94 (1985).
- Figge, J., Webster, T., Smith, T. F. & Paucha, E. *J. Virol.* **62**, 1814-1818 (1988).
- Rustgi, A. K., Dyson, N. & Bernards, R. *Nature* **352**, 541-544 (1991).
- Blackwood, E. M. & Eisenman, R. N. *Science* **251**, 1211-1217 (1991).
- Vousden, K. & Jat, P. *Oncogene* **4**, 153-158 (1989).
- Kaczmarek, L., Hyland, J. K., Watt, R., Rosenberg, M. & Baserga, R. *Science* **228**, 1313-1315 (1985).
- Stabel, S., Argos, P. & Philipson, L. *EMBO J.* **4**, 2329-2336 (1988).
- Chellappan, S. P., Hiebert, S., Mudry, M., Horowitz, J. M. & Nevins, J. R. *Cell* **65**, 1053-1061 (1991).
- Wagner, S. & Green, M. R. *Nature* **352**, 189-190 (1991).
- Whyte, P., Williamson, N. M. & Harlow, E. *Cell* **56**, 67-75 (1989).
- Jones, N. C., Richter, J. D., Weeks, D. L. & Smith, L. D. *Molec. cell. Biol.* **3**, 2131-2142 (1983).
- Symonds, G. et al. *Oncogene* **4**, 285-294 (1989).
- Ramsay, G., Evan, G. I. & Bishop, J. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7742-7746 (1984).
- Harlow, E., Franza, B. R. Jr & Schley, C. J. *J. Virol.* **55**, 533-546 (1985).
- Alitalo, K. et al. *Nature* **306**, 274-277 (1983).

ACKNOWLEDGEMENTS. I thank J. M. Bishop (UCSF), in whose laboratory most of this work was performed, and D. Dina (Chiron) for support and encouragement, R. Bernards for communicating results before publication, N. Jones, A. Brusk, M. Scott and W. Lee for gifts of plasmids, G. Ramsay and E. Harlow for antibodies, and K. Thudium for technical assistance.

CORRECTIONS

Plasma fireballs formed by microwave interference in air

Y. H. Ohtsuki & H. Ofuruton

Nature **350**, 139-141 (1991).

THE fifth line in the right-hand column on page 139 states: "A standing wave should be set up in the cavity by interference between a propagating wave and a reflected wave, and in this case there should be theoretically six antinodes at which the electric field is strongest." However, because we could not identify the mode which existed in the cavity and the guide wavelength must be used in the calculation, there should be theoretically plural (at most five) antinodes in the cavity. This correction does not affect any of the conclusions in the paper.

A protein-tyrosine phosphatase with sequence similarity to the SH2 domain of the protein-tyrosine kinases

Shi-Hsiang Shen, Lison Bastien, Barry I. Posner & Pierre Chrétien

Nature **352**, 736-739 (1991).

WE have discovered errors made in the assembly of our published cDNA sequence affecting the end of the coding region in Fig. 1 of the above paper. One -g- should be inserted between nucleotides 2,028 and 2,029. Thus, the open reading frame of the PTP1C gene terminates at nucleotide 2,046 instead of nucleotide 2,082. The corresponding amino acid sequence of the last 6 residues at the C-terminus of PTP1C from nucleotides 2,029 to 2,046 is GSLKRR. The correct sequence is available on the EMBL database under accession number X62055. These errors do not affect the conclusions of the paper in any way.

ERRATUM

Preparation and structure of the alkali-metal fulleride, A₄C₆₀

R. M. Fleming, M. J. Rosseinsky, A. P. Ramirez, D. W. Murphy, J. C. Tully, R. C. Haddon, T. Siegrist, R. Tycko, S. H. Giarum, P. Marsh, G. Dabbagh, S. M. Zahurak, A. V. Makhija & C. Hampton

Nature **352**, 701-703 (1991).

THE positions of the alkali atoms, originally reported as (0, $\frac{1}{2}$, z) with z = 0.28, should read (x, $\frac{1}{2}$, 0) with x = 0.22. These are the positions used in the calculations for Table 1 and the positions illustrated in Fig. 2.

Instrumentation market focus

Once again London plays host to Labweek '91 (5-7 November), an annual exhibition featuring new product lines in the laboratory instrumentation arena. Highlights include an electroporation system with dual-voltage selection, software programs for HPLC simulations and a fully automated capillary zone electrophoresis system.

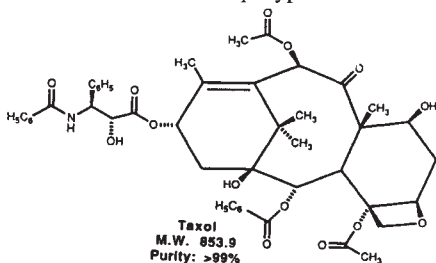
AS part of a new working relationship between Skatron Instruments and BTX Biotechnology, Skatron is distributing the new BTX Electro Cell Manipulator 600 electroporation system (*Reader Service No. 100*). Special design features of



ECM 600 — new transfection tool from BTX.

the ECM 600 include a choice of two voltage ranges to suit all applications (2.5 kV/500 V), additive capacitor bank resistance timing and instantaneous monitoring — all in a compact unit, Skatron says. At the turn of a single control, users can switch settings from bacterial transformation to mammalian or plant transfection. Internal monitoring devices measure and display true peak voltage and pulse lengths after every discharge. Recharging takes less than five seconds, the company says. Sterile reaction cuvettes are available in two sizes (compatible with other systems). The system is supported by access to a list of protocols and the Electronic Genetics Data Base operated by BTX. Stop by Skatron's exhibit on stand 762.

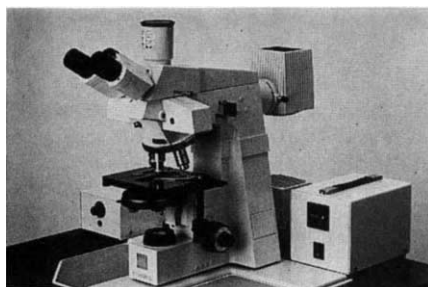
Taxol, an **anti-tumour and anti-leukaemic agent that is isolated from the bark of the Pacific yew tree, *Taxus brevifolia***, is now available from Calbiochem (*Reader Service No. 101*). It has been shown to promote the assembly of microtubules and inhibit tubulin disassembly. Taxol has also been reported to act on tumour necrosis factor (TNF) receptors and TNF release, and *in vitro* studies demonstrate polyploidization in



Tap into Calbiochem's taxol supply.

taxol-resistant human leukaemia cells. Calbiochem says that long-term effects of single injections of taxol upon peripheral nerve axons have also been shown. Taxol costs DM210 (Germany) for a 5-mg quantity. Catch Calbiochem on stand 653.

Labweek '91 will be the launching pad for the new Axioskop FS, a fixed-stage version of the Axioskop **pyramid microscope** from Carl Zeiss (*Reader Service No. 102*). Instead of the stage moving towards the fixed objective to allow focusing, as in the Axioskop microscope, the FS model is focused by moving the upper part of the stand towards the stage. Smooth movement of the microscope tube is ensured by incorporating a counterbalance mechanism into the stand, Zeiss says. Specimens with a maxi-



Fixed-stage Axioskop FS microscope.

mum height of 42 mm can be accommodated. A $\times 5$ DIC nosepiece, which accepts the full range of Zeiss ICS infinity optics for brightfield, phase contrast, DIC and fluorescence imaging techniques, is a standard feature. Zeiss also offers a range of TV, photo-documentation and micromanipulation accessories for the Axioskop FS. A separate power supply provides either 12 V/50 W or 12 V/100 W halogen illumination. See Zeiss on stand 706.

Dry runs

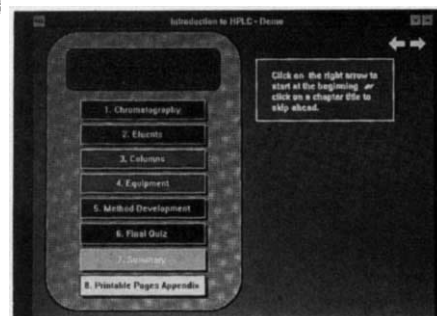
DryLab G/plus is a new software program from LC Resources that allows the chromatographer to **simulate gradient HPLC separations** on the computer for rapid and accurate methods development (*Reader Service No. 103*). Following the entry of retention-time data from two experimental runs, the user examines the effects of changes in gradient time, initial and final mobile-phase composition, column dimensions, flow rate and size of the packing particle by viewing chroma-



DryLab G/plus — chromatographic methods development by computer simulation.

tograms simulated for these conditions. Both linear and multi-segmented gradients can be generated in seconds, the company says. DryLab G/plus adjusts gradient slope automatically to maintain constant selectivity while the user determines the effects of changes in column dimensions or flow rate. The program is useful for simulating biological separations. As so few actual runs are required, sample consumption is minimized, LC Resources says. And, changes in retention time, which can occur as a result of repeated injections of biological samples onto the column, are also minimized. The program can be run on any IBM PC, XT AT or compatible computer. A demonstration disk is available on request. LC Resources will be on stand 205.

"Introduction to high-performance liquid chromatography" is an **animated software training program** from Auto-scribe, which is designed to run under Microsoft Windows on a PC (*Reader Service No. 104*). Operation of chromatographic components, eluent flow



Windows-based HPLC teaching program.

through a system and other subject areas are demonstrated, with action sequences in full colour. The interactive program also features an "electronic" chromatograph, which shows the resulting changes in a chromatogram when various experimental parameters are altered. Quizzes provide a means of on-screen self-testing. With an organization akin to a book, the various "chapter" headings include: chromatography (equilibria, retention, resolution); eluents (properties, selectivity, gradient elution); columns (normal and reversed phase, ion exchange, size exclusion); equipment (pumps, gradient controllers, sampling valves, columns, detectors, data systems); and method development (qualitative, quantitative and preparative separations). A demonstration disk is available on request. Autoscribe also offers video training programmes on a variety of topics including centrifugation, protein electrophoresis, and health and safety. Number 205 is the Autoscribe stand.

Molecular medley

Camlab is distributing Amresco's Pro-PAGE ready-to-use **denaturing acrylamide gel mix** for discontinuous gel electrophoresis of proteins (*Reader Service No. 105*). Pro-PAGE is supplied pre-mixed with Tris and 0.1% SDS for denaturing electrophoresis. The resolving gel is available in 6%, 8% and 12% acrylamide concentrations (acrylamide:bis-acrylamide ratio is 37.5:1) and provided with a 4% stacking-gel mix for discontinuous systems. Use of Pro-PAGE requires no pre-measuring or mixing, just add the polymerizing agents (ammonium persulphate and TEMED) and pour, Camlab says. All three concentrations of Pro-PAGE are sold in either 5 × 100-ml or 4 × 500-ml quantities, costing £68.40 and £255.60 (UK), respectively. Peruse Camlab's stand, number 608.

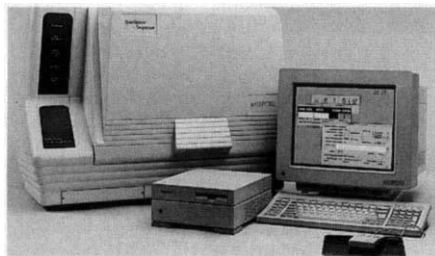
Fully automated operation and precise temperature control designed to ensure run-to-run reproducibility are features of the new BioFocus 3000 **capillary zone electrophoresis system** from Bio-Rad (*Reader Service No. 106*). The system can



Automated operation and temperature control — features of BioRad's BioFocus 3000 CZE system.

be used to perform a variety of methods — capillary zone electrophoresis, isoelectric focusing (IEF), non-gel sieving and micellar electrokinetic capillary chromatography. Capillary IEF with BioFocus 3000, made possible because of the patented coating, can be used to resolve samples with a pI difference of as little as 0.02 pH unit, Bio-Rad says. The system's fast-scanning UV/visible detector can monitor from one to 20 wavelengths simultaneously. Three-dimensional plots of absorption spectra versus migration time can be used to confirm peak purity and aid in peak identification. Temperature control of the capillary and the autosampler are independent. The ability to cool samples in the autosampler allows unattended overnight running of labile samples. Additional design features of the BioFocus 3000 include 32 vial positions on both the inlet and outlet ends of the capillary, a fraction collection capability for the recovery of separated components, and multiple sample injection techniques (pressure displacement and electromigration). The system software operates in the Windows 3.0 environment and the data-base structure of the software makes it easy to archive/retrieve both the results of a run and the run parameters. See stand 700.

The new BaseStation DNA sequencer from Millipore automates the three key components in the sequencing process, namely electrophoresis, band detection and base identification (*Reader Service No. 107*). The instrument reads, processes and displays the sequence data as four-lane sequence ladders. At the heart



Millipore's BaseStation DNA sequencer.

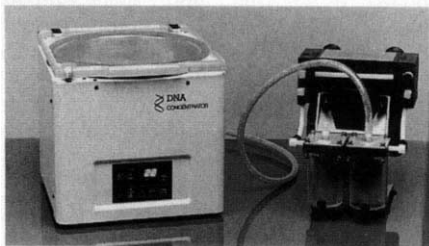
of the system is the charge-coupled device camera, which contains an array of hundreds of thousands of individual detectors. Each detector can register a single photon of light and can distinguish between 2¹⁶ shades of gray. This, Millipore says, results in an accurate and sensitive image of the band patterns. Bases are identified in real time, as electrophoresis progresses, resulting in run times of about six hours and the possibility of making two runs in a day. For high throughput, 20 samples can be analysed in a single run. The BaseStation uses a single fluorescent primer in the sequencing reactions. Included with the BaseStation is a Sun computer and icon-based software. Sequences can be edited

and stored to tape, or exported to external data-base programs for homology searches. Millipore will be showing off its wares on stand 710.

With the recent announcement of version 1.0 of MacDNASIS Pro for Macintosh computers, Hitachi's **DNA and protein sequence analysis software** will be available shortly on both Macintosh and PC computer systems (*Reader Service No. 108*). MacDNASIS Pro 1.0 offers the following: sequence input and editing, primary and secondary structure prediction, "shotgun" fragment connection, restriction mapping, and similarity search and retrieval of genetic information from GenBank, EMBL, NBRF-PIR and SWISS-PROT data bases supplied on CD-ROM. Additional features include a PCR primer design function and a multiple sequence alignment function. The \$2,450 (US) MacDNASIS Pro package for Macs will be available in early 1992. See stand 424.

Concentrated efforts

Although only the size of a microfuge, the new **mini centrifugal evaporator** from V. A. Howe has a capacity of up to 30 microtubes (*Reader Service No. 109*).



V. A. Howe's purpose-built DNA concentrator has a 30-microtube capacity.

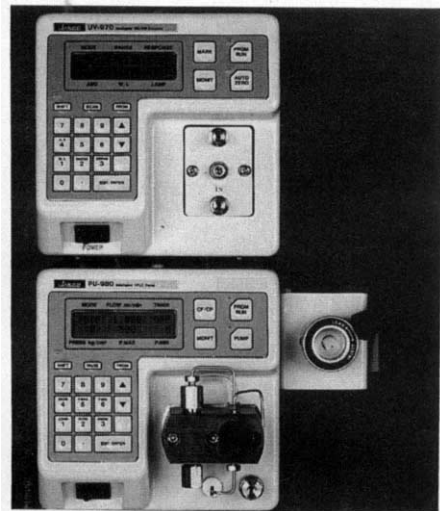
The DNA concentrator offers full temperature control up to 69°C. As temperature of the bowl is displayed at all times, the user is able to gauge the recovery of precious amounts of genetic material from small volumes of solvents with little or no sample loss, the company says. The DNA concentrator can be used with a diaphragm pump, eliminating the need for a solvent trap. The pump can be controlled directly by the evaporator, which, the company adds, prevents "vacuum jump" and additional sample loss. V. A. Howe will be on stand 804.

On display for the first time at Lab-week '91 will be Techne's **sample concentrator unit**, which is designed for use with the company's DB-3 range of Dri-Block heaters (*Reader Service No. 110*). The unit can be used to accelerate the concentration of a large number of samples in minutes, Techne says. The latest Dri-Block heaters with an extended temperature range to 200°C and models with pre-sets and digital displays offer flexibility and a convenient means of tempera-

ture control. For procedures requiring a temperature stability of up to $\pm 0.005^\circ\text{C}$, users can choose from one of Techne's water baths. Techne will be on stand 718.

Analytical instrumentation

Jasco will be launching the new LC-900 range of **HPLC pumps and detectors** on Ciba Corning's stand number 416 at Labweek '91 (Reader Service No. 111).



Pump it up — see stand 416.

The tapered flow cell of the UV-975 variable-wavelength detector provides reduced refractive index effects, the company says. The fast-scan mode allows on-the-fly scanning without the need for stop-flow or bypass mechanisms. This, coupled with an extensive programming facility (64 programme steps/10 files) increases the flexibility of the instrument still further. To complement the UV-975 detector, Jasco offers the PU-980 "intelligent" pump, which has built-in system controller functions for complex ternary low-pressure and binary high-pressure gradients. External event outputs and the six-way switching valve can also be actuated.

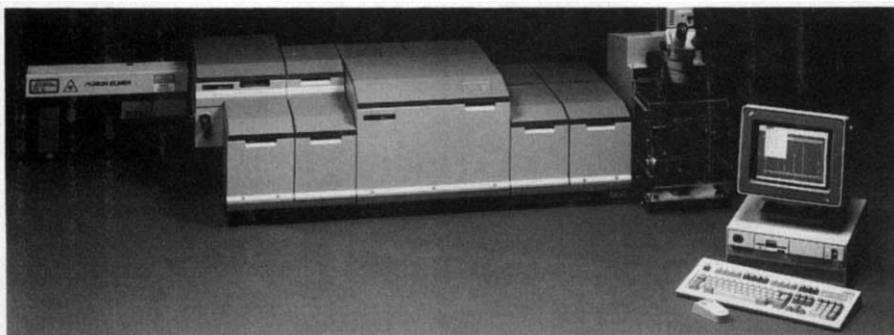
To meet the changing needs of analytical and research chemists, Perkin-Elmer has introduced the System 2000, a new concept in **FT-IR spectroscopy** based upon an Open Architecture design in three key areas — sampling, optical

configuration and data handling (Reader Service No. 112). The conventional internal and external sample compartments have been discarded in favour of an interferometer module with up to four equivalent output beams arranged symmetrically at the four corners. Each output beam can be configured with a sampling station, including single- and double-sample compartments, FT-IR microscopes, GC-IR interface, TG-IR interface and near-IR FT Raman. This set-up, while meeting current sampling requirements, leaves the system open for future upgrades and reconfiguration, the company says. In addition, the system's optical flexibility allows multiple sources, beamsplitters and detectors to be fitted to the same unit to cover all or any part of the range $15,000$ to 20 cm^{-1} . Open data handling means the System 2000 can be used with an industry-standard 386-based computer platform. See stand 404.

Bench-top aids

Take the tedium out of Soxhlet extractions with the new **EMEA stirring electromantle unit** from Electrothermal Engineering (Reader Service No. 113). Based on the earlier EME electromantle controlled-heating unit for Soxhlet extraction, the new EMEA unit is equipped with an in-built stirring facility that is continuously variable from 50 to 1,000 r.p.m. The addition of a stirring capability results in a significant shortening of the reflux cycle time and faster extractions, the company says. The EMEA electromantle consists of a bank of three or six heating recesses in a choice of sizes to hold 100-, 250-, 500- or 1,000-ml flasks. Each recess features individually-controlled heating up to 450°C . The three-bank unit features a single stirring control; the six-bank version, two controls. With user-safety in mind, solid-state energy regulators have been built into every heating recess to prevent sparking caused by mechanical switching. Additionally, a pre-formed stainless steel safety screen covers every heater. See stand 472.

Unicam has added three new **bench-top pH and ion-selective meters** to its electrochemistry product range (Reader

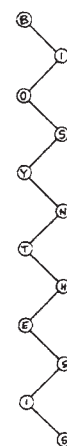


Open sampling architecture — the hallmark of Perkin-Elmer's FT-IR System 2000.

ADVERTISEMENTS

PEPTIDES

Around the World



- Two weeks turnaround
- \$65.00/residue, 200 umole scale
- Amino acid analysis included
- Experience and reliability you can count on

Other Services:
Carrier Conjugation
Antiserum Production
Protein-Peptide Sequencing

1-800-227-0627

Your Source of Peptides

BIO-SYNTHESIS

Fax 1-214-420-0442

Reader Service No.18

LIPOSOMES

REPRODUCIBLE — HOMOGENEOUS —
EXTREMELY RAPID



Our new LIPOSOMAT guarantees rapid (5ml within 60 minutes) preparation of uniformly sized unilamellar liposomes. Liposome size is selectable between 25 and ca. 600 nm diameter. Sample volume is normally 5-10ml or up to 200ml if used in combination with an additional instrument. Lipid concentration can be up to 300 mg/ml.

DIANDROM

POB 126, D-8000 Munich 65, FRG

Reader Service No.17

Synthetic Oligonucleotides

Custom made probes and primers,
synthesised, purified and shipped in 3 days

£3.50 per base

(All inclusive)

0.2 μM scale (approx. 500 μg)

Modified bases, biotin or digoxigenin labelling,
data base analysis and full design facilities

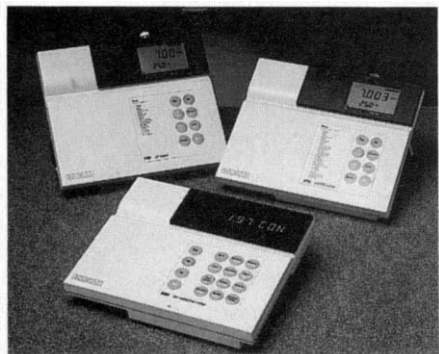
Molecular Medicine Unit,
King's College School of Medicine & Dentistry,
Denmark Hill, London SE5 8RX

Tel. XX-44-(0)71 - 326 3126

FAX XX-44-(0)71 - 733 3877

Reader Service No.56

Service No. 114). For making routine pH measurements, Unicam recommends its 9450 model. For those researchers requiring a research-grade instrument with a resolution of 0.001 pH, in addition to a direct concentration readout for routine ion-selective analysis, Unicam suggests the 9455 model. The 9460 top-

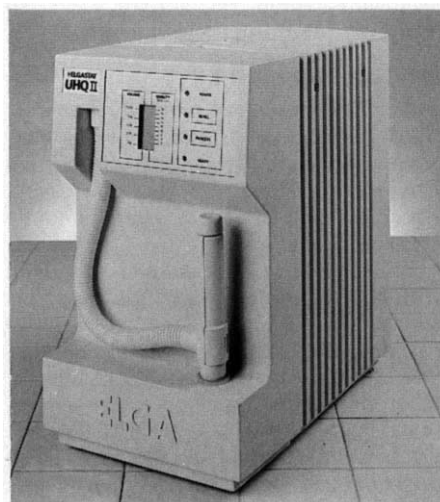


Unicam's meters measure up.

of-the-range model is not only designed to provide fast and reliable ion-selective analysis, but can also perform incremental techniques such as multipoint calibration, dual-electrode input, electrode identification, stability display and data logging. All three models offer a choice of resolution (for pH), simulta-

neous temperature display, autocalibration with a selection of three or five recognised buffers and an RS232 serial output. Unicam will be on stand 410.

The Elgastat UHQII **water purification unit** from Elga is designed to have wide application in areas of research, such as tissue culture, which demand water of the highest standards of purity (*Reader Service No.115*). Each UHQII unit is a complete purification system that comes in a compact bench-top unit. The ultra-pure water produced by the UHQII has a conductivity of $>18 \text{ M}\Omega \text{ cm}^{-1}$ at 25°C , a TOC content of $<20 \text{ p.p.b.}$ and a bacterial content below 1 CFU ml^{-1} , Elga says. The inclusion of a UV-disinfection step at the final treatment stage produces water that is suitable for making up culture media supple-



Ultra-pure water on tap.

ments and reagents, and for PCR procedures. Elga will be on stand 212.

Handy gadgets

Is there too much clutter and computer paraphernalia on your desk top? Radleys thinks it may have the answer. The company's new **mouse house** is a stand



Reclaim your desk top with Radley's mouse house.

made of acrylic that is designed to hold the computer keyboard above the desk top and inclined towards the operator at an angle of 15° (*Reader Service No. 116*). This set-up should provide users with a clutter-free work surface with enough space to manipulate the mouse. The 6-mm thick, heavy-duty acrylic is designed to hold even the weightiest keyboard. The keyboard itself is held in place by a 15-mm high lip on the front edge of the stand. Six non-slip rubber feet prevent slippage of the stand on the desk top. Radley's mouse house is available in two sizes. An acrylic book stand is an optional extra. Radley's range of laboratory accessories will be featured on stand 305.

The storage and disposal of hazardous radioactive waste is a problem common to most research laboratories. It need not be a headache, Stuart Scientific says, with the introduction of the new SW2 range of **safety bins** (*Reader Service No. 117*). Three models are available — all constructed of 10-mm thick Perspex and designed to withstand and protect users from beta-emitting isotopes up to ^{32}P , Stuart says. The table-top version, SW2/3, comes with a lift-off lid and is



Radioactive waste bin on wheels.

small enough to sit on the bench, in a fume cupboard or freezer, but sufficiently roomy to hold comfortably a CB1-size "Cinbin". Stuart also offers two floor-standing bins, the SW2/4 and SW2/5, each featuring a (foot) pedal-operated lid, a handle and a label pocket. To facilitate movement of the bin around the laboratory, both floor models are mounted on castors. As an additional safety measure, the bins are supplied with a polythene inner container. See Stuart Scientific on stand 811. □

These notes are compiled by Diane Gershon from information provided by the manufacturers. To obtain further details, use the reader service card bound inside the journal. Prices quoted are sometimes nominal, and apply only within the country indicated.

ADVERTISEMENTS

SINGER
INSTRUMENTS



COMPLETE SYSTEMS FOR
XENOPUS
OOCYTE INJECTION

INT. SALES: +44. 984. 40226.

FAX: + 44. 984. 41166

Singer Instrument Co. Ltd.

Roadwater, Watchet, Somerset,
England, TA23 0QL. Tel: Washford (0984) 40226.
Telex: 337492 Comcab G

Reader Service No.50

Custom DNA
Purified & Shipped in 24 hours
Worldwide services

100 μg minimum • \$7.50 per base
Call to obtain our new catalog

Research Genetics

U.S. & Canada 1-800-533-4363

• U.K. 0-800-89-1393 •

• FAX 1-205-536-9016 •

Reader Service No.36

Prescriptions for pharmaceutical research

In a time of recession, pharmaceutical companies remain committed to research and development. But the commercial pressures they face mean recruits fresh from their PhDs may be in for a culture shock.

In August when Bayer described its restructuring plans to financial analysts in the City of London, the German pharmaceutical company was careful to point out that none of its rationalization would result in any sort of cutback in research and development (R&D).

That research funding should be sacrosanct in a business which spends a higher proportion of its turnover on R&D than any other industry apart from aerospace, (see figure), is ample demonstration of how highly it is valued. Any serious international contender puts 15 per cent of its turnover into research. It is estimated that more than \$10,000 million per annum is invested in pharmaceutical research worldwide.

The United States, Japan, France, Italy, Germany, Switzerland and Britain account for about 80 per cent of this.

The top spending companies in 1990 were Bristol-Myers Squibb at \$657 million, Glaxo \$654 million, Hoechst \$613 million, Bayer \$487 million, Sandoz, \$484 million and Johnson and Johnson \$419 million. Overall, R&D expenditure has risen by 15 per cent annually since 1983.

All of which translates into a vast number of jobs for scientists. For example, Smith-Kline Beecham has 5,000 R&D staff, Rhône-Poulenc Rorer 3,400 and Glaxo 6,000.

The Alexandria-based American Association of Pharmaceutical Scientists, which functions as a professional society (tracking salaries, organizing scientific meetings and so on) has 5,500 members (few of them MDs), but reckons that its potential membership is 15,000. This is probably a good estimate of the total number of scientists active in research in US industrial laboratories.

Most companies recruit PhDs for their basic research efforts. Indeed this is one

business where first-degree graduates are treated as technicians rather than high flyers. Alex Bearn, the now retired director of research at Merck (who is a Trustee of the Rockefeller University and an emeritus professor there) says that the most noticeable change during his spell in industry was the trend to the recruitment of physicians as researchers. He says that the pharmaceutical industry is in transition from a chemistry-based industry to one centred on human biology, made possible because of the increased number of physicians with a background in research. Yet the skilled organic chemist remains dominant throughout the industry.

The cliché goes that pharmaceutical companies are so sensitive to the needs of scientists that they work hard to foster the right environment for creative thought, putting together small highly focused teams, and leaving the researchers to get on with it.

But times are changing, according to Alain Schreiber, senior vice president for R&D at the French-American multinational Rhône-Poulenc Rorer. "The days are gone when sales and marketing would wait for R&D to come up with something and then go out and sell it."

The drugs companies now aim to make their R&D as market oriented as possible. This means using medical intelligence and strategic planning, and deciding which research projects get priority. The aim is to balance high risk with low risk, short with medium term, and the costs and duration of clinical research with the state of the market.

In other words, pharmaceutical companies want to know what an idea is worth at the time one of their research teams has it. They are now looking from the end of the process before starting their development, and asking questions such as, How much money will we have to spend on promotion? Who are the competitors? How are we going to develop it? What are the clinical end points?

The pressure to do this stems not just from growing competition but from the fact that the drugs companies claim to have seen the return on R&D spending declining, with the number of new drug launches decreasing 10 per cent each year from 1969 to 1985.

According to the US Pharmaceutical Manufacturers Association the cost of developing a new drug has risen from an average of \$125 million in 1987 to \$230

million in 1990, although this is misleading because it is an average which includes the cost of the many research projects that come to nothing.

So it seems increasingly that researchers in the pharmaceutical industry will not be able to hide with their test tubes. Decision-making is being pushed down to R&D, and researchers are being made more aware of the commercial constraints of the drugs business.

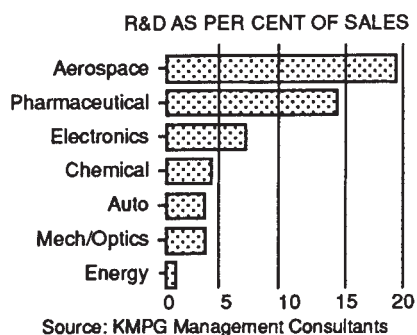
Newly qualified PhDs joining a pharmaceutical company are bound to find it a shock to the system, according to Schreiber. In a university lab someone may have concentrated on making one molecule in the most sophisticated way possible. They are then put in a company lab and told to make as many molecules as they can by the simplest routes.

New recruits can expect to be paid around £16,000-£20,000 in Britain, depending on discipline. Medicinal chemists are in particularly short supply, whereas molecular biologists and pharmacologists are easier to find.

The American Association of Pharmaceutical Scientists is conscious of skill-shortages throughout its field, many of which have developed only in the past few years. (Companies such as Merck each recruit between 200 and 300 scientists a year for their research laboratories). Numerically, the shortage of chemists is most apparent, but there are fields such as statistics in which the pharmaceutical companies must fight hard to win their share from the relatively small output of the US universities.

US entry salaries are around \$45,000-\$50,000, but there is a premium on those with scarce skills, who can expect to earn \$50,000-\$55,000. One measure of the shortage of skilled people in the United States is that even new joiners are often given a money present, on top of their salary, but industry spokesmen are coy about the sums involved. A Merck spokesman says "Oh, No! Not that Much!" when asked if it is as much as a year's salary, but gives the impression that a newly qualified PhD could expect a substantial fraction of his or her starting salary.

More senior academics look for bigger prizes when they move from university to industrial laboratories. The big companies are used to offering productive researchers not just a much higher salary but employment for their existing research group as well.



THE TOP 15 PHARMACEUTICAL COMPANIES WORLDWIDE, 1989-90

	Company	Pharmaceutical sales	
		(\$million)	% change
1 (1)	Merck & Co (US)	5,405.5	+11.0
2 (2)	Glaxo (UK)	4,679.5	+2.2
3 (—)	Bristol-Myers Squibb (US)	4,442.0	+11.4
4 (4)	Bayer (Germany)	4,237.8	+14.1
5 (3)	Hoechst (Germany)	4,200.5	+6.1
6 (14)	Eastman Kodak (US)	4,009.0	+8.6
7 (5)	Ciba-Geigy (Switzerland)	3,775.9	+6.9
8 (—)	SmithKline Beecham (UK/US)	3,668.8	+7.8
9 (8)	Sandoz (Switzerland)	3,464.1	+10.1
10 (7)	Am Home Products (US)	3,276.5	+3.3
11 (6)	Takeda (Japan)	3,237.2	-6.7
12 (9)	Eli Lilly (US)	2,934.0	+16.0
13 (10)	Abbott (US)	2,785.0	+7.2
14 (12)	Warner-Lambert (US)	2,694.0	+7.4
15 (11)	Pfizer (US)	2,687.0	+5.8

Ranking in brackets are those for the previous year.
Source: Scrip's Pharmaceutical Company League Tables.

A newly qualified PhD who succeeds in adding value to the team may end up as team leader. The next career step is to become section head, and then department head. The next promotion would be to research director. Department heads can expect to earn £50,000 plus benefits in Britain; research directors at least £100,000. The tendency is to move from company to company to get these promotions.

Scientists can also move from company to company chasing the projects that interest them. A researcher who is any good can afford to be more interested in science than in his or her employer.

At certain levels of seniority, companies will decide they want to recruit a scientist from another company, rather than picking someone from a university. This is usually because they need someone who has experienced a commercial environment. Companies also recognize that it is good to get new blood, bringing in ideas and intelligence from their competitors.

But they would run away from any suggestion that they would go out and recruit to get a competitive position by luring teams away from other companies. As Schreiber puts it, "It's difficult to put your finger on the expert that counts, and so you would probably end up trying to lure away 25 people".

Pharmaceuticals is one industry where the scientists do not have to get stuck in the back room, and can in fact get to be managing director or chief executive. In a way, this is the best indicator there is of what counts most in these companies. But although companies are increasingly trying to tempt people out of the lab and into commercial roles they also recognise that some people want to be scientists and nothing else.

Pay and promotion is organized to provide lateral mobility which allows people who stay in R&D to reach the same point on the salary scale as peers who move beyond the lab. The British

company ICI Pharmaceuticals, for example, pays scientists who are good enough up to £50,000 to stay at the laboratory bench; Rhône-Poulenc Rorer offers those who opt to stay in the laboratory the same stock options and variable compensation as researchers who move up the management hierarchy. A senior researcher will be paid between \$90,000 and \$100,000.

Most European companies recruit R&D staff on an international basis, whereas they will tend to stick to their domestic markets for sales and marketing and other managers. Roughly a third of the scientists at Ciba-Geigy's research labs in Basle, Switzerland are from overseas, and among recent recruits the proportion of other nationalities has risen to 70 per cent.

Dr Max Wilhelm, head of research and development at the company's Basle pharmaceutical division recognizes that Ciba-Geigy has to look outside Switzerland to find certain skills. For example he says that Britain is the best source of research-oriented physicians for conducting clinical trials, and of pharmacologists, while Schreiber points out that Britain is one of the few countries that still trains physiologists — a discipline that has gone out of fashion in the United States.

However, few companies have international pay and conditions structures, and so salary levels vary according to national markets. In the UK for example, salaries are generally lower than in France or Germany. In any case, scientists joining pharmaceutical companies can expect to earn 20 per cent more than their counterparts in universities.

Although recruitment consultants report more pan-European recruitment, they say that the preferred route to moving country is to join a company and get an overseas posting.

To meet local regulatory requirements all big drug companies do research in more than one country, and the prospect is that a scientist joining a pharmaceutical

company will be posted from one country to another. It is also recognized that the only way to build a multinational business is to move people around. But companies also need to choose carefully who they move, because it is an expensive business, costing around 2½ times an individual's annual salary.

Another cliché about the pharmaceutical industry is that it offers good job security because it is recession proof — after all, people are always going to be ill, and buying drugs is not the kind of purchase that can be deferred.

This has in part proved true during the current recession, with the drugs companies turning in rising profits while other science and technology based industries such as aerospace and electronics have seen profits fall.

But although the money keeps rolling in, there is a vast amount of restructuring going on in the industry. The pharmaceutical business is unusual for a global industry in being highly fragmented, with the top 25 companies accounting for only 50 per cent of the market. There were estimated to be \$45,000 million worth of mergers and acquisitions between 1988 and 1990, and as the table shows, these mergers resulted in a notable reshuffle among the top 15 companies.

Although, as in the case of Bayer, R&D is looked on as sacrosanct, restructuring is resulting in change in this side of the business. For example, Ciba-Geigy had R&D projects in 18 therapeutic areas in 1980, and cut this to seven by 1989; when the US company Rorer merged with the drugs arm of the French company Rhône-Poulenc in 1990 there were 70 projects in the pipeline which had to be rationalised and put in order of priority. More mergers are likely in future because only high turnover companies can fund the R&D programmes, inevitably causing upheavals for their research scientists.

One of the other frustrations of pharmaceutical research is the huge number of dead ends. Project attrition rates are frightening — typically for every 10,000 compounds patented, animal trials are conducted on only 1,000, human trials on only 100, and only one will actually gain marketing approval. There are so many false trails that people may worry that they will be chastised for not doing a good job. Schreiber says that one of the most important bits of psychology in nurturing a research team is not to punish people for failing.

The aim of pharmaceutical company managers is to produce a perfect consensus between science and marketing. While this may force scientists to adopt a different outlook from that they have grown up with in universities there is no doubt there will continue to be plenty of opportunities for those who can stomach the change.

Nuala Moran